GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:18:16 ; Search time 1242.59 Seconds
(without alignments)
13865.275 Million cell updates/sec

Title: US-09-905-114-1
Perfect score: 592
Sequence: 1 gagaagctaaacggaaattc.....gtctgaatcgaaggccctgg 592

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	476.8	80.5	1023	6	AX329742	AX329742 Sequence
	2	476.8	80.5	1023	6	AX332286	AX332286 Sequence
	3	476.8	80.5	1023	6	AX332789	AX332789 Sequence
	4	476.8	80.5	1023	6	AX410746	AX410746 Sequence
	5	476.8	80.5	1023	9	HSU56814	U56814 Human DNase
	6	476.8	80.5	1067	9	BC015831	BC015831 Homo sapi
	7	476.8	80.5	1079	9	AF047354	AF047354 Homo sapi
	8	475.2	80.3	1108	9	HSU75744	U75744 Homo sapien
	9	459.2	77.6	1331	10	AF039852	AF039852 Rattus no
	10	459.2	77.6	1417	10	RNU75689	U75689 Rattus norv
	11	452.8	76.5	1208	6	AR047845	AR047845 Sequence
	12	452.8	76.5	1208	6	E11687	E11687 cDNA encodi
	13	435.2	73.5	941	10	MMU76110	U76110 Mus musculu
	14	435.2	73.5	1124	10	AF047355	AF047355 Mus muscu
	15	433.6	73.2	2110	10	BC012671	BC012671 Mus muscu
	16	387.2	65.4	1223	6	AX400016	AX400016 Sequence
	17	280.8	47.4	1250	5	AF059612	AF059612 Xenopus l
	18	173.2	29.3	1295	4	BTPANDNAI	AJ001538 Bos tauru
	19	172.2	29.1	868	4	AF311922	AF311922 Bos tauru
c	20	171.4	29.0	392	6	AX330074	AX330074 Sequence
c	21	171.4	29.0	392	6	AX407983	AX407983 Sequence
	22	163.2	27.6	1210	5	AB038776	AB038776 Rana cate
	23	160.6	27.1	1283	5	AB041732	AB041732 Cynops py
	24	154	26.0	1161	4	D82875	D82875 Oryctolagus
	25	152.2	25.7	1312	10	BC014718	BC014718 Mus muscu
	26	152.2	25.7	1337	10	BC030394	BC030394 Mus muscu
	27	151	25.5	1297	4	AB048832	AB048832 Sus scrof
	28	150.8	25.5	1156	10	RNDESOXI	X56060 Rat mRNA fo
	29	150.8	25.5	1184	10	AF397151	AF397151 Rattus no
	30	150.8	25.5	1313	10	AF397150	AF397150 Rattus no
	31	150.6	25.4	1180	10	D83038	D83038 Mus musculu
	32	149.4	25.2	1039	6	AX030575	AX030575 Sequence
c	33	149.4	25.2	1039	6	AX030591	AX030591 Sequence
	34	149.4	25.2	1039	6	AX268671	AX268671 Sequence
	35	149.4	25.2	1039	9	HUMDNASEI	M55983 Human DNase
	36	149.4	25.2	1149	9	BC029437	BC029437 Homo sapi
	37	149.4	25.2	2054	9	HSA298844	AJ298844 Homo sapi
	38	148.4	25.1	1066	5	AB013751	AB013751 Gallus ga
	39	147.8	25.0	783	6	AX268673	AX268673 Sequence
	40	147.8	25.0	858	6	AX268675	AX268675 Sequence
	41	147.8	25.0	1039	6	AR019468	AR019468 Sequence
	42	147.8	25.0	1548	6	AX268721	AX268721 Sequence
	43	147.8	25.0	1548	6	AX268722	AX268722 Sequence
	44	147.8	25.0	1554	6	AX268701	AX268701 Sequence
	45	147.8	25.0	1554	6	AX268702	AX268702 Sequence

ALIGNMENTS

RESULT 1
AX329742
LOCUS AX329742 1023 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 251 from Patent WO0194629.
ACCESSION AX329742
VERSION AX329742.1 GI:18102720
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature

gene sets
JOURNAL Patent: WO 0194629-A 251 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1. .1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 312 a 244 c 240 g 227 t
ORIGIN

Query Match 80.5%; Score 476.8; DB 6; Length 1023;
Best Local Similarity 87.8%; Pred. No. 1.7e-119;
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
Db     241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
Db     301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
Db     361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
Db     421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
Db     481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
Db     541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
Db     601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
Db     661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
Db     721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
Db     781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832
```

RESULT 2
AX332286
LOCUS AX332286 1023 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 2795 from Patent WO0194629.
ACCESSION AX332286
VERSION AX332286.1 GI:18122920
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 2795 13-DEC-2001;
Avalon Pharmaceuticals (US)

FEATURES Location/Qualifiers
source 1. .1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 312 a 244 c 240 g 227 t
ORIGIN

Query Match 80.5%; Score 476.8; DB 6; Length 1023;
Best Local Similarity 87.8%; Pred. No. 1.7e-119;
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
Db     241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
Db     301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
Db     361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
Db     421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
Db     481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
Db     541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
Db     601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
Db     661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
Db     721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
Db     781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832
```

RESULT 3
AX332789
LOCUS AX332789 1023 bp DNA linear PAT 10-JAN-2002
DEFINITION Sequence 3298 from Patent WO0194629.
ACCESSION AX332789
VERSION AX332789.1 GI:18123423
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 3298 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1. .1023
/organism="Homo sapiens"

```
                     /db_xref="taxon:9606"
BASE COUNT      312 a    244 c    240 g    227 t
ORIGIN

  Query Match           80.5%;  Score 476.8;  DB 6;  Length 1023;
  Best Local Similarity  87.8%;  Pred. No. 1.7e-119;
  Matches  520;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          ||||||||  |||  ||||||||||  |  |||||||||  ||||||||||||||||||||||||
Db    241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||||||||||||||  ||||||||||||||  ||  ||||||||  ||||||
Db    301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          ||  ||    ||  ||  |  ||||||  ||||||||||  ||||||||||||||||||||||||
Db    361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          ||  ||||||||||||||||||||  ||  |||  ||||  ||||||||  ||||||||||||  ||
Db    421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||||  |||||  ||||||  |||  |||  ||||||
Db    481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          ||  ||  |||||||  ||||||||  ||  ||||||||||||||||||||||||||||||||
Db    541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||||  |||||||  |||||||  |||||||  |
Db    601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          ||  ||||||||||||||||||||||||||||||||||||||  ||||||||||  |||||  ||
Db    661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          ||||||||||  ||||||||||||||||  |  ||  ||||  ||||  ||  |||  |  ||||||
Db    721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            ||  |||||  ||||||||||||||||  |  ||  |||||  |  |||||||||
Db    781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832


RESULT 4
AX410746
LOCUS       AX410746               1023 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 3393 from Patent WO0229103.
ACCESSION   AX410746
VERSION     AX410746.1  GI:21443451
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 3393 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1. .1023
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. U56814"
BASE COUNT      312 a    244 c    240 g    227 t
ORIGIN
```

```
  Query Match           80.5%;  Score 476.8;  DB 6;  Length 1023;
  Best Local Similarity  87.8%;  Pred. No. 1.7e-119;
  Matches  520;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          ||||||||  |||  ||||||||||  |  |||||||||  ||||||||||||||||||||||||
Db    241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||||||||||||||  ||||||||||||||  ||  ||||||||  ||||||
Db    301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          ||  ||    ||  ||  |  ||||||  ||||||||||  ||||||||||||||||||||||||
Db    361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          ||  ||||||||||||||||||||  ||  |||  ||||  ||||||||  ||||||||||||  ||
Db    421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||||  |||||  ||||||  |||  |||  ||||||
Db    481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          ||  ||  |||||||  ||||||||  ||  ||||||||||||||||||||||||||||||||
Db    541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||||  |||||||  |||||||  |||||||  |
Db    601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          ||  ||||||||||||||||||||||||||||||||||||||  ||||||||||  |||||  ||
Db    661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          ||||||||||  ||||||||||||||||  |  ||  ||||  ||||  ||  |||  |  ||||||
Db    721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            ||  |||||  ||||||||||||||||  |  ||  |||||  |  |||||||||
Db    781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832


RESULT 5
HSU56814
LOCUS       HSU56814               1023 bp    mRNA    linear   PRI 24-JUL-1997
DEFINITION  Human DNase1-Like III protein (DNAS1L3) mRNA, complete cds.
ACCESSION   U56814
VERSION     U56814.1  GI:1399718
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1023)
  AUTHORS   Rodriguez,A.M., Rodin,D., Nomura,H., Morton,C.C., Weremowicz,S. and
            Schneider,M.C.
  TITLE     Identification, localization, and expression of two novel human
            genes similar to deoxyribonuclease I
  JOURNAL   Genomics 42 (3), 507-513 (1997)
  MEDLINE   97349121
   PUBMED   9205125
REFERENCE   2  (bases 1 to 1023)
  AUTHORS   Schneider,M.C. and Rodriguez,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-1996) M.C. Schneider, Renal Division, Brigham and
            Women's Hospital, 75 Francis Street, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1. .1023
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /clone_lib="Stratagene liver (#937224)"
                     /note="corresponds to EST clones 82269 (GenBank Accession
                     Number T68985, T69063), 82738 (GenBank Accession Number
                     T73558, T73653), 78422 (GenBank Accession Number T61400,
                     T61368)"
     gene            1. .1023
                     /gene="DNAS1L3"
     CDS             25. .942
                     /gene="DNAS1L3"
                     /codon_start=1
                     /product="DNaseI-Like III protein"
                     /protein_id="AAB63967.1"
                     /db_xref="GI:1399719"
                     /translation="MSRELAPLLLLLLSIHSALAMRICSFNVRSFGESKQEDKNAMDV
                     IVKVIKRCDIILVMEIKDSNNRICPILMEKLNRNSRRGITYNYVISSRLGRNTYKEQY
                     AFLYKEKLVSVKRSYHYHDYQDGDADVFSREPFVVWFQSPHTAVKDFVIIPLHTTPET
                     SVKEIDELVEVYTDVKHRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPRFVWLIG
                     DQEDTTVKKSTNCAYDRIVLRGQEIVSSVVPKSNSVFDFQKAYKLTEEEALDVSDHFP
                     VEFKLQSSRAFTNSKKSVTLRKKTKSKRS"
BASE COUNT      312 a    244 c    240 g    227 t
ORIGIN

  Query Match             80.5%;  Score 476.8;  DB 9;  Length 1023;
  Best Local Similarity   87.8%;  Pred. No. 1.7e-119;
  Matches  520;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          ||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||||
Db     241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db     301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db     361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          |||||||||||||||| ||||||||||||  ||||| |||||| ||| ||| ||||||
Db     481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db     541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          ||||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db     601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || ||||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db     661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||||| | || ||||  |||| || |||  | ||||||
Db     721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
             || ||||| |||||||||||||||| | || |||||  | |||||||||
Db     781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832


RESULT 6
BC015831
LOCUS       BC015831               1067 bp    mRNA    linear   PRI 29-OCT-2001
DEFINITION  Homo sapiens, deoxyribonuclease I-like 3, clone MGC:27146
            IMAGE:4723725, mRNA, complete cds.
ACCESSION   BC015831
VERSION     BC015831.1  GI:16198370
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1067)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
            Greene, Mark Ketteman and Anuradha Madan

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 38 Row: p Column: 7
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 4826697.
FEATURES             Location/Qualifiers
     source          1. .1067
                     /organism="Homo sapiens"
                     /db_xref="LocusID:1776"
                     /db_xref="taxon:9606"
                     /clone="MGC:27146 IMAGE:4723725"
                     /tissue_type="Liver"
                     /clone_lib="NIH_MGC_76"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     CDS             68. .985
                     /codon_start=1
                     /product="deoxyribonuclease I-like 3"
                     /protein_id="AAH15831.1"
                     /db_xref="GI:16198371"
                     /translation="MSRELAPLLLLLLSIHSALAMRICSFNVRSFGESKQEDKNAMDV
                     IVKVIKRCDIILVMEIKDSNNRICPILMEKLNRNSRRGITYNYVISSRLGRNTYKEQY
                     AFLYKEKLVSVKRSYHYHDYQDGDADVFSREPFVVWFQSPHTAVKDFVIIPLHTTPET
                     SVKEIDELVEVYTDVKHRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPRFVWLIG
                     DQEDTTVKKSTNCAYDRIVLRGQEIVSSVVPKSNSVFDFQKAYKLTEEEALDVSDHFP
                     VEFKLQSSRAFTNSKKSVTLRKKTKSKRS"
BASE COUNT      327 a    256 c    249 g    235 t
ORIGIN

  Query Match             80.5%;  Score 476.8;  DB 9;  Length 1067;
  Best Local Similarity   87.8%;  Pred. No. 1.7e-119;
  Matches  520;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          ||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||||
Db     284 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 343

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db     344 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 403

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db     404 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 463

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     464 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 523
```

```
Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||| |||||||||||||  ||||| |||||| ||| ||| ||||||
Db    524 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 583

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||||
Db    584 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 643

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||||||| |||||||| ||||||| |||||||
Db    644 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 703

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || ||||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db    704 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 763

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||| | || ||||  |||| || |||  | ||||||
Db    764 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 823

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
          || ||||| |||||||||||||||| | || |||||  | |||||||||
Db    824 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 875


RESULT 7
AF047354
LOCUS       AF047354               1079 bp    mRNA    linear   PRI 16-SEP-1998
DEFINITION  Homo sapiens liver and spleen DNase precursor (LSD) mRNA, complete
            cds.
ACCESSION   AF047354
VERSION     AF047354.1  GI:2905785
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Baron,W.F., Pan,C.Q., Spencer,S.A., Ryan,A.M., Lazarus,R.A. and
            Baker,K.P.
  TITLE     Cloning and characterization of an actin-resistant DNase I-like
            endonuclease secreted by macrophages
  JOURNAL   Gene 215 (2), 291-301 (1998)
  MEDLINE   98382522
   PUBMED   9714828
REFERENCE   2  (bases 1 to 1079)
  AUTHORS   Baron,W.F., Pan,C.Q., Spencer,S.A., Ryan,A.M., Lazarus,R.A. and
            Baker,K.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-1998) Molecular Biology, Genentech Inc., 1 DNA
            Way, South San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .1079
                     /gene="LSD"
     CDS             71. .988
                     /gene="LSD"
                     /note="actin resistant acivity; macrophage specific DNase
                     I; LS-DNase"
                     /codon_start=1
                     /product="liver and spleen DNase precursor"
                     /protein_id="AAC35752.1"
                     /db_xref="GI:2905786"
                     /translation="MSRELAPLLLLLLSIHSALAMRICSFNVRSFGESKQEDKNAMDV
                     IVKVIKRCDIILVMEIKDSNNRICPILMEKLNRNSRRGITYNYVISSRLGRNTYKEQY
                     AFLYKEKLVSVKRSYHYHDYQDGDADVFSREPFVVWFQSPHTAVKDFVIIPLHTTPET
                     SVKEIDELVEVYTDVKHRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPRFVWLIG
                     DQEDTTVKKSTNCAYDRIVLRGQEIVSSVVPKSNSVFDFQKAYKLTEEEALDVSDHFP
                     VEFKLQSSRAFTNSKKSVTLRKKTKSKRS"
     sig_peptide     71. .130
                     /gene="LSD"
                     /evidence=experimental
     mat_peptide     131. .985
                     /gene="LSD"
                     /product="liver and spleen DNase"
BASE COUNT      330 a    260 c    252 g    237 t
ORIGIN

  Query Match           80.5%;  Score 476.8;  DB 9;  Length 1079;
  Best Local Similarity  87.8%;  Pred. No. 1.7e-119;
  Matches  520;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||| |||||||||||||||||||||||
Db    287 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 346

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db    347 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 406

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db    407 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 466

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db    467 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 526

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          |||||||||||||||| |||||||||||  ||||| |||||| ||| ||| ||||||
Db    527 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 586

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||||
Db    587 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 646

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||||||| |||||||| ||||||| |||||||
Db    647 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 706

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db    707 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 766

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||| | || ||||  |||| || |||  | ||||||
Db    767 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 826

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
          || ||||| |||||||||||||||| | || |||||  | |||||||||
Db    827 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 878


RESULT 8
HSU75744
LOCUS       HSU75744               1108 bp    mRNA    linear   PRI 17-JUN-1998
DEFINITION  Homo sapiens DNase gamma mRNA, complete cds.
ACCESSION   U75744
VERSION     U75744.1  GI:3236319
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1108)
  AUTHORS   Shiokawa,D., Hirai,M. and Tanuma,S.
  TITLE     cDNA cloning of human DNase gamma: chromosomal localization of its
            gene and enzymatic properties of recombinant protein
  JOURNAL   Apoptosis 3 (2), 89-95 (1998)
REFERENCE   2  (bases 1 to 1108)
  AUTHORS   Shiokawa,D. and Tanuma,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-1996) Biochemistry, Science Univ. of Tokyo,
```

```
          Shinjuku-ku Ichigaya Funagawaracho, Tokyo 162, Japan
FEATURES             Location/Qualifiers
     source          1. .1108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             93. .1010
                     /function="apoptosis"
                     /function="deoxyribonuclease"
                     /codon_start=1
                     /product="DNase gamma"
                     /protein_id="AAC23652.1"
                     /db_xref="GI:3236320"
                     /translation="MSRELAPLLLLLLSIHSALAMRICSFNVRSFGESKQEDKNAMDV
                     IVKVIKRCDIILVMEIKDSNNRICPILMEKLNRNSRRGITYNYVISSRLGRKTYKEQY
                     AFLYKEKLVSVKRSYHYHDYQDGDADVFSREPFVVWFQSPHTAVKDFVIIPLHTTPET
                     SVKEIDELVEVYTDVKHRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPRFVWLIG
                     DQEDTTVKKSTNCAYDRIVLRGQEIVSSVVPKSNSVFDFQKAYKLTEEEALDVSDHFP
                     VEFKLQSSRAFTNSKKSVTLRKKTKSKRS"
BASE COUNT      329 a    261 c    271 g    247 t
ORIGIN
```

Query Match 80.3%; Score 475.2; DB 9; Length 1108;
Best Local Similarity 87.7%; Pred. No. 4.7e-119;
Matches 519; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

```
Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||
Db     309 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 368

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||| |||||||||||||| |||||||||||||| || |||||||| ||||||
Db     369 CTTGGAAGAAAAACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 428

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| ||||||||| ||||||||||||||||||||||||
Db     429 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 488

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     489 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 548

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          |||||||||||||||| |||||||||||||   ||||| |||||| ||| ||| ||||||
Db     549 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 608

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||||
Db     609 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 668

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||||||||||| ||||||| ||||||| |||
Db     669 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 728

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db     729 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 788

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||| | || ||||  |||| || |||  | ||||||
Db     789 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 848

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            || |||||  |||||||||||||||| | || |||||  | |||||||||
Db     849 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 900
```

```
RESULT 9
AF039852
LOCUS       AF039852                1331 bp    mRNA    linear   ROD 07-AUG-1998
DEFINITION  Rattus norvegicus DNAseY mRNA, complete cds.
ACCESSION   AF039852
VERSION     AF039852.1  GI:3395771
KEYWORDS    .
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1331)
  AUTHORS   Liu,Q.Y., Pandey,S., Singh,R.K., Lin,W., Ribecco,M.,
            Borowy-Borowski,H., Smith,B., LeBlanc,J., Walker,P.R. and
            Sikorska,M.
  TITLE     DNaseY: a rat DNaseI-like gene coding for a constitutively
            expressed chromatin-bound endonuclease
  JOURNAL   Biochemistry 37 (28), 10134-10143 (1998)
  MEDLINE   98332537
   PUBMED   9665719
REFERENCE   2  (bases 1 to 1331)
  AUTHORS   Liu,Q.Y., Singh,R.K., Lin,W. and Sikorska,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-DEC-1997) Institute for Biological Sciences, National
            Research Council, 1200 Montreal Road, Bldg. M-54, Box 4, Ottawa, ON
            K1A 0R6, Canada
FEATURES             Location/Qualifiers
     source          1. .1331
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /tissue_type="liver"
     CDS             153. .1085
                     /function="apoptotic endonuclease"
                     /note="similar to rat DNAse I; chromatin-bound"
                     /codon_start=1
                     /product="DNAseY"
                     /protein_id="AAC28937.1"
                     /db_xref="GI:3395772"
                     /translation="MSLYPASPYLASLLLFILALHGALSLRLCSFNVRSFGESKKENH
                     NAMDIIVKIIKRCDLILLMEIKDSNNNICPMLMEKLNGNSRRSTTYNYVISSRLGRNT
                     YKEQYAFLYKEKLVSVKAKYLYHDYQDGDTDVFSREPFVVWFQAPFTAAKDFVIVPLH
                     TTPETSVKEIDELADVYTDVRRRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPNF
                     VWLIGDQEDTTVKKSTSCAYDRIVLRGQEIVNSVVPRSSGVFDFQKAYELSEEEALDV
                     SDHFPVEFKLQSSRAFTNSRKSVSLKKKKKGSRS"
BASE COUNT      351 a    333 c    320 g    327 t
ORIGIN
```

Query Match 77.6%; Score 459.2; DB 10; Length 1331;
Best Local Similarity 86.0%; Pred. No. 1.1e-114;
Matches 509; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```
Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| || ||||| ||| ||| | |||  ||||||||||| ||||||||||||||
Db     384 GAGAAGCTGAATGGAAACTCACGAAGAAGCACGACATACAACTACGTGATTAGCTCTCGG 443

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||||||||||||||||||| ||||| || || ||||| ||||||
Db     444 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTCCTCTACAAGGAGAAGCTGGTGTCT 503

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||   ||  |||||||||| |||||||||| ||||||| |||| ||||||||||||
Db     504 GTGAAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 563

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||| |||||||| | |||| ||| |||| ||||||||||||||||||||||
Db     564 GAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCACTGCTGCCAAGGACTTCGTGATTGTC 623

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||| ||||||| || ||||| || ||||||| |||||| |||||||||||||| ||||||
Db     624 CCCTTGCACACAACTCCTGAAACCTCCGTTAAAGAGATAGATGAGCTGGCTGACGTCTAC 683

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || ||||||| | | || ||||| |||||||||||||| |||||||||| |||||||||
Db     684 ACGGATGTGAGAAGACGATGGAAGGCAGAGAATTTCATCTTCATGGGTGATTTCAATGCT 743

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||||||||| |||||||  ||||||| ||||||||
Db     744 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGTTTGAGGACAGACCCCAAC 803
```

```
Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||||| |||||||||||||||||||||||||||||||||| | ||| |||
Db    804 TTTGTTTGGCTGATTGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACCAGCTGTGCC 863

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||||| |||||| | |||||| | || ||||||||||  | ||| |  || |
Db    864 TATGACAGGATTGTGCTTCGCGGACAAGAGATAGTCAACTCTGTGGTTCCCCGCTCCAGT 923

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            ||||||||| || ||||||||||   ||||||||||  | |||||||||
Db    924 GGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGG 975
```

RESULT 10
RNU75689

```
LOCUS       RNU75689                1417 bp    mRNA    linear   ROD 15-JUN-1998
DEFINITION  Rattus norvegicus DNase gamma mRNA, complete cds.
ACCESSION   U75689
VERSION     U75689.1  GI:3220024
KEYWORDS    .
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1417)
  AUTHORS   Shiokawa,D., Ohyama,H., Yamada,T., Takahashi,K. and Tanuma,S.
  TITLE     Identification of an endonuclease responsible for apoptosis in rat
            thymocytes
  JOURNAL   Eur. J. Biochem. 226 (1), 23-30 (1994)
  MEDLINE   95045594
   PUBMED   7957253
REFERENCE   2  (bases 1 to 1417)
  AUTHORS   Shiokawa,D. and Tanuma,S.
  TITLE     Molecular cloning and expression of a cDNA encoding an apoptotic
            endonuclease DNase gamma
  JOURNAL   Biochem. J. 332 (Pt 3), 713-720 (1998)
  MEDLINE   98285539
   PUBMED   9620874
REFERENCE   3  (bases 1 to 1417)
  AUTHORS   Shiokawa,D. and Tanuma,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-1996) Biochemistry, Science Univ. of Tokyo,
            Shinjuku-ku ichigaya funagawaracho, Tokyo 162, Japan
FEATURES             Location/Qualifiers
     source          1. .1417
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /tissue_type="spleen"
     CDS             203. .1135
                     /function="apoptosis"
                     /note="deoxyribonuclease"
                     /codon_start=1
                     /product="DNase gamma"
                     /protein_id="AAC40134.1"
                     /db_xref="GI:3220025"
                     /translation="MSLYPASPYLASLLLFILALHGALSLRLCSFNVRSFGESKKENH
                     NAMDIIVKIIKRCDLILLMEIKDSNNNICPMLMEKLNGNSRRSTTYNYVISSRLGRNT
                     YKEQYAFLYKEKLVSVKAKYLYHDYQDGDTDVFSREPFVVWFQAPFTAAKDFVIVPLH
                     TTPETSVKEIDELADVYTDVRRRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPNF
                     VWLIGDQEDTTVKKSTSCAYDRIVLRGQEIVNSVVPRSSGVFDFQKAYELSEEEALDV
                     SDHFPVEFKLQSSRAFTNSRKSVSLKKKKKGSRS"
BASE COUNT      386 a    353 c    333 g    345 t
ORIGIN
```

```
  Query Match             77.6%;  Score 459.2;  DB 10;  Length 1417;
  Best Local Similarity   86.0%;  Pred. No. 1.1e-114;
  Matches  509;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          ||||||||  || ||||| ||| ||| | |||  |||||||||||  ||||||||||||||
Db    434 GAGAAGCTGAATGGAAACTCACGAAGAAGCACGACATACAACTACGTGATTAGCTCTCGG 493

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||||||||||||||||||||||||||||||||| ||||| || || ||||| ||||||
Db    494 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTCCTCTACAAGGAGAAGCTGGTGTCT 553

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||   ||  |||||||||||| |||||||||| ||||||| |||| ||||||||||||
Db    554 GTGAAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 613

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||| ||||||||| | |||| ||| |||| ||||||||||||||||||||
Db    614 GAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCACTGCTGCCAAGGACTTCGTGATTGTC 673

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||| ||||||| || ||||| || ||||||| |||||| |||||||||||||| ||||||
Db    674 CCCTTGCACACAACTCCTGAAACCTCCGTTAAAGAGATAGATGAGCTGGCTGACGTCTAC 733

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || ||||||| | | || ||||| |||||||||||||| |||||||||| |||||||||
Db    734 ACGGATGTGAGAAGACGATGGAAGGCAGAGAATTTCATCTTCATGGGTGATTTCAATGCT 793

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||| ||||||| ||||||| |||||||
Db    794 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGTTTGAGGACAGACCCCAAC 853

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||| ||||||||||||||||||||||||||||||||| | ||| |||
Db    854 TTTGTTTGGCTGATTGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACCAGCTGTGCC 913

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||||| |||||| | |||||| | || ||||||||||  | ||| |  || |
Db    914 TATGACAGGATTGTGCTTCGCGGACAAGAGATAGTCAACTCTGTGGTTCCCCGCTCCAGT 973

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            ||||||||| || ||||||||||   ||||||||||  | |||||||||
Db    974 GGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGG 1025
```

RESULT 11
AR047845

```
LOCUS       AR047845                1208 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 2 from patent US 5821103.
ACCESSION   AR047845
VERSION     AR047845.1  GI:5970188
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1208)
  AUTHORS   Tanuma,S.-i.
  TITLE     Deoxyribonuclease
  JOURNAL   Patent: US 5821103-A 2 13-OCT-1998;
FEATURES             Location/Qualifiers
     source          1. .1208
                     /organism="unknown"
BASE COUNT      319 a    304 c    283 g    302 t
ORIGIN
```

```
  Query Match             76.5%;  Score 452.8;  DB 6;  Length 1208;
  Best Local Similarity   85.3%;  Pred. No. 6.4e-113;
  Matches  505;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          ||||||||  || ||||| ||| ||| | |||  |||||||||||  ||||||||||||||
Db    243 GAGAAGCTGAATGGAAACTCACGAAGAAGCACGACATACAACTACGTGATTAGCTCTCGG 302

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||||||||||||||||||||||||||||||||| ||||| || || ||||| ||||||
Db    303 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTCCTCTACAAGGAGAAGCTGGTGTCT 362

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||   ||  |||||||||||| |||||||||| ||||||| |||| ||||||||||||
```

```
Db     363 GTGAAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 422

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240

Db     423 GAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCACTGCTGCCAAGGACTTCGTGATTGTC 482

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300

Db     483 CCCTTGCACACAACTCCTGAAACCTCCGTTAAAGAGATAGATGAGCTGGCTGACGTCTAC 542

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360

Db     543 ACGGATGTTAGAAGACGATGGAAGGCAGAGATTTTCATCTTCATGGGTGATTTCAATGCT 602

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420

Db     603 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGTTTGAGGACAGACCCCAAC 662

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480

Db     663 TTTGTTTGGCTGATTGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACCAGCTGTGCC 722

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540

Db     723 TATGACAGGATTGTCGTTCGCGGACAAGAGATAGTCAACTCTGTGGTTCCCCGCTCCAGT 782

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592

Db     783 GGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGG 834
```

```
RESULT 12
E11687
LOCUS       E11687                  1208 bp    RNA     linear   PAT 29-SEP-1997
DEFINITION  cDNA encoding novel deoxyribonuclease(DNase) gamma which cut
            specifically linker site of chromatin DNA.
ACCESSION   E11687
VERSION     E11687.1  GI:22025323
KEYWORDS    JP 1996187079-A/1.
SOURCE      Rattus rattus.
  ORGANISM  Rattus rattus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1208)
  AUTHORS   Tanuma,Y.
  TITLE     NEW DEOXYRIBONUCLEASE
  JOURNAL   Patent: JP 1996187079-A 1 23-JUL-1996;
            TANUMA YASUKAZU
COMMENT     OS   Rattus rattus (rat)
            PN   JP 1996187079-A/1
            PD   23-JUL-1996
            PF   06-SEP-1995 JP 1995255647
            PR   06-SEP-1994 JP 94P   239518
            PI   TANUMA YASUKAZU
            PC   C12N9/16,C07K14/47,C07K16/40,C12N1/21,C12N15/09,(C12N1/21, PC
             C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .1208
            FT                   /organism='Rattus rattus'
            FT                   /tissue_type='Thymus, Spleen' FT
               /cell_type='Cell nucleus'
            FT   5'UTR           1. .11
            FT   CDS             12. .944
            FT                   /product='Novel deoxynbonuclease(DNase) gamma
            FT                   which cut
            FT                   specifically linker site of chromatin DNA' FT
             precursor_RNA   12. .86
            FT   3'UTR           945. .1208.
FEATURES            Location/Qualifiers
     source         1. .1208
                    /organism="Rattus rattus"
                    /db_xref="taxon:10117"
BASE COUNT      319 a    304 c    283 g    302 t
ORIGIN
```

```
  Query Match           76.5%;  Score 452.8;  DB 6;  Length 1208;
  Best Local Similarity  85.3%;  Pred. No. 6.4e-113;
  Matches  505;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60

Db     243 GAGAAGCTGAATGGAAACTCACGAAGAAGCACGACATACAACTACGTGATTAGCTCTCGG 302

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120

Db     303 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTCCTCTACAAGGAGAAGCTGGTGTCT 362

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180

Db     363 GTGAAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 422

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240

Db     423 GAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCACTGCTGCCAAGGACTTCGTGATTGTC 482

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300

Db     483 CCCTTGCACACAACTCCTGAAACCTCCGTTAAAGAGATAGATGAGCTGGCTGACGTCTAC 542

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360

Db     543 ACGGATGTTAGAAGACGATGGAAGGCAGAGATTTTCATCTTCATGGGTGATTTCAATGCT 602

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420

Db     603 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGTTTGAGGACAGACCCCAAC 662

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480

Db     663 TTTGTTTGGCTGATTGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACCAGCTGTGCC 722

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540

Db     723 TATGACAGGATTGTCGTTCGCGGACAAGAGATAGTCAACTCTGTGGTTCCCCGCTCCAGT 782

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592

Db     783 GGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGG 834
```

```
RESULT 13
MMU76110
LOCUS       MMU76110                941 bp    mRNA    linear   ROD 26-JAN-1999
DEFINITION  Mus musculus DNase gamma mRNA, complete cds.
ACCESSION   U76110
VERSION     U76110.1  GI:4098207
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 941)
  AUTHORS   Shiokawa,D., Hatanaka,T. and Tanuma,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-1996) Biochemistry, Science Univ. of Tokyo,
            Shinjuku-ku ichigaya funagawaracho, Tokyo 162, Japan
FEATURES            Location/Qualifiers
     source         1. .941
                    /organism="Mus musculus"
                    /strain="C57 black"
```

```
                     /db_xref="taxon:10090"
     CDS             7. .939
                     /function="apoptosis"
                     /note="deoxyribonuclease"
                     /codon_start=1
                     /product="DNase gamma"
                     /protein_id="AAD09222.1"
                     /db_xref="GI:4098208"
                     /translation="MSLHPASPRLASLLLFILALHDTLALRLCSFNVRSFGASKKENH
                     EAMDIIVKIIKRCDLILLMEIKDSSNNICPMLMEKLNGNSRRSTTYNYVISSRLGRNT
                     YKEQYAFVYKEKLVSVKTKYHYHDYQDGDTDVFSREPFVVWFHSPFTAVKDFVIVPLH
                     TTPETSVKEIDELVDVYTDVRSQWKTENFIFMGDFNAGCSYVPKKAWQNIRLRTDPKF
                     VWLIGDQEDTTVKKSTSCAYDRIVLCGQEIVNSVVPRSSGVFDFQKAYDLSEEEALDV
                     SDHFPVEFKLQSSRAFTNNRKSVSLKKRKKGNRS"
BASE COUNT      255 a    229 c    228 g    229 t
ORIGIN

  Query Match           73.5%;  Score 435.2;  DB 10;  Length 941;
  Best Local Similarity  83.4%;  Pred. No. 4.1e-108;
  Matches  494;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| || ||||||||| ||| | ||| |||||||||||||||||||||| |||||
Db    238 GAGAAGCTGAATGGAAATTCACGAAGAAGCACAACATACAACTATGTGATTAGTTCTCGA 297

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||| || ||||| |||||||||| |||| || || ||||| ||||||
Db    298 CTTGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGCTGGTGTCT 357

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||   ||  |||| |||||| |||||||||| ||||||| |||| ||||||||||||
Db    358 GTGAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 417

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||| |||||||| || ||||  || ||||||||||||||||||||||||||
Db    418 GAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCGTGATTGTC 477

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||| ||||||| || || ||||| ||||||| |||||| |||||||||| |||||||||
Db    478 CCCTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCGATGTCTAC 537

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || ||||||| | | |  ||||| |||||||||||| |||||||||| ||||| ||
Db    538 ACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATTTCAACGCC 597

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          ||||| ||||| |||||||||||||||||| || |||| ||  ||||||||||||||||||
Db    598 GGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGAGGACGGACCCCAAG 657

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||||| |||||||||||||||||| |||||||||||||| || | ||| |||
Db    658 TTTGTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCAGCTGTGCC 717

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||| | |||||| | || |||||||| |  | ||| |  || |
Db    718 TATGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTTCCCCGTTCCAGT 777

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
          |||||||||| || ||||||||||    ||||||||    | |||||||||
Db    778 GGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 829


RESULT 14
AF047355
LOCUS       AF047355               1124 bp    mRNA    linear   ROD 16-SEP-1998
DEFINITION  Mus musculus liver and spleen DNase precursor (LSD) mRNA, complete
            cds.
ACCESSION   AF047355
VERSION     AF047355.1  GI:2905787
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1124)
  AUTHORS   Baron,W.F., Pan,C.Q., Spencer,S.A., Ryan,A.M., Lazarus,R.A. and
            Baker,K.P.
  TITLE     Cloning and characterization of an actin-resistant DNase I-like
            endonuclease secreted by macrophages
  JOURNAL   Gene 215 (2), 291-301 (1998)
  MEDLINE   98382522
   PUBMED   9714828
REFERENCE   2  (bases 1 to 1124)
  AUTHORS   Baron,W.F., Pan,C.Q., Spencer,S.A., Ryan,A.M., Lazarus,R.A. and
            Baker,K.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-1998) Molecular Biology, Genentech Inc., 1 DNA
            Way, South San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1. .1124
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     gene            1. .1124
                     /gene="LSD"
     CDS             173. .1105
                     /gene="LSD"
                     /note="actin resistant acivity; macrophage specific DNase
                     I; LS-DNase"
                     /codon_start=1
                     /product="liver and spleen DNase precursor"
                     /protein_id="AAC35753.1"
                     /db_xref="GI:2905788"
                     /translation="MSLHPASPRLASLLLFILALHDTLALRLCSFNVRSFGASKKENH
                     EAMDIIVKIIKRCDLILLMEIKDSSNNICPMLMEKLNGNSRRSTTYNYVISSRLGRNT
                     YKEQYAFVYKEKLVSVKTKYHYHDYQDGDTDVFSREPFVVWFHSPFTAVKDFVIVPLH
                     TTPETSVKEIDELVDVYTDVRSQWKTENFIFMGDFNAGCSYVPKKAWQNIRLRTDPKF
                     VWLIGDQEDTTVKKSTSCAYDRIVLCGQEIVNSVVPRSSGVFDFQKAYDLSEEEALDV
                     SDHFPVEFKLQSSRAFTNNRKSVSLKKRKKGNRS"
     sig_peptide     173. .247
                     /gene="LSD"
                     /evidence=not_experimental
     mat_peptide     248. .1102
                     /gene="LSD"
                     /product="liver and spleen DNase"
BASE COUNT      299 a    273 c    273 g    279 t
ORIGIN

  Query Match           73.5%;  Score 435.2;  DB 10;  Length 1124;
  Best Local Similarity  83.4%;  Pred. No. 4.1e-108;
  Matches  494;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| || ||||||||| ||| | ||| |||||||||||||||||||||| |||||
Db    404 GAGAAGCTGAATGGAAATTCACGAAGAAGCACAACATACAACTATGTGATTAGTTCTCGA 463

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||| || ||||| |||||||||| |||| || || ||||| ||||||
Db    464 CTTGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGCTGGTGTCT 523

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||   ||  |||| |||||| |||||||||| ||||||| |||| ||||||||||||
Db    524 GTGAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 583

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||| |||||||| || ||||  || ||||||||||||||||||||||||||
Db    584 GAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCGTGATTGTC 643

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||| ||||||| || || ||||| ||||||| |||||| |||||||||| |||||||||
Db    644 CCCTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCGATGTCTAC 703

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || ||||||| | | |  ||||| |||||||||||| |||||||||| ||||| ||
Db    704 ACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATTTCAACGCC 763

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
```

```
          ||||| ||||| |||||||||||||||||| || |||| ||  ||||||||||||||||
Db    764 GGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGAGGACGGACCCCAAG 823

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||| |||||||||||||||||| |||||||||||||| || | ||| |||
Db    824 TTTGTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCAGCTGTGCC 883

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||| | |||||| | || |||||||| |  | ||| |  || |
Db    884 TATGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTTCCCCGTTCCAGT 943

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            ||||||||| || ||||||||||    ||||||||   | |||||||||
Db    944 GGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 995


RESULT 15
BC012671
LOCUS       BC012671               2110 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, Similar to deoxyribonuclease 1-like 3, clone
            MGC:13854 IMAGE:4160709, mRNA, complete cds.
ACCESSION   BC012671
VERSION     BC012671.1  GI:15215118
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2110)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 18 Row: i Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 6681204.
FEATURES             Location/Qualifiers
     source          1. .2110
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="MGC:13854 IMAGE:4160709"
                     /tissue_type="Liver, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             151. .1083
                     /codon_start=1
                     /product="Similar to deoxyribonuclease 1-like 3"
                     /protein_id="AAH12671.1"
                     /db_xref="GI:15215119"
                     /translation="MSLHPASPRLASLLLFILALHDTLALRLCSFNVRSFGASKKENH
                     EAMDIIVKIIKRCDLILLMEIKDSSNNICPMLMEKLNGNSRRSTIYNYVISSRLGRNT
                     YKEQYAFVYKEKLVSVKTKYHYHDYQDGDTDVFSREPFVVWFHSPFTAVKDFVIVPLH
                     TTPETSVKEIDELVDVYTDVRSQWKTENFIFMGDFNAGCSYVPKKAWQNIRLRTDPKF
                     VWLIGDQEDTTVKKSTSCAYDRIVLCGQEIVNSVVPRSSGVFDFQKAYDLSEEEALDV
                     SDHFPVEFKLQSSRAFTNNRKSVSLKKRKKGNRS"
```

```
BASE COUNT      526 a    581 c    464 g    539 t
ORIGIN

  Query Match           73.2%;  Score 433.6;  DB 10;  Length 2110;
  Best Local Similarity  83.3%;  Pred. No. 1.2e-107;
  Matches  493;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| || ||||||||| ||| | ||| || |||||||||||||||||| |||||
Db    382 GAGAAGCTGAATGGAAATTCACGAAGAAGCACAATATACAACTATGTGATTAGTTCTCGA 441

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||| || ||||| |||||||||| |||| || || ||||| ||||||
Db    442 CTTGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGCTGGTGTCT 501

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||   || |||| |||||| |||||||||| ||||||| |||| ||||||||||||
Db    502 GTGAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 561

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||| |||||||| || ||||  || ||||||||||||||||||||||||
Db    562 GAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCGTGATTGTC 621

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||| ||||||| || || ||||| ||||||| |||||| |||||||||  |||||||||
Db    622 CCCTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCGATGTCTAC 681

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || ||||||| | | |  ||||| |||||||||||| |||||||||| ||||| ||
Db    682 ACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATTTCAACGCC 741

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          ||||| ||||| |||||||||||||||||| || |||| ||  ||||||||||||||||
Db    742 GGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGAGGACGGACCCCAAG 801

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||| |||||||||||||||||| |||||||||||||| || | ||| |||
Db    802 TTTGTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCAGCTGTGCC 861

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||| | |||||| | || |||||||| |  | ||| |  || |
Db    862 TATGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTTCCCCGTTCCAGT 921

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            |||||||||| || ||||||||||    ||||||||   | |||||||||
Db    922 GGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 973


Search completed: December  9, 2002, 11:30:44
Job time : 1260.59 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 08:59:14 ; Search time 170.616 Seconds
(without alignments)
7813.939 Million cell updates/sec

Title: US-09-905-114-1
Perfect score: 592
Sequence: 1 gagaagctaaacggaaattc.....gtctgaatcgaaggccctgg 592

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	592	100.0	592	24	AAD29088	Bovine fertility a
	2	592	100.0	671	24	AAD29089	Bovine 22kDà recom
	3	539	91.0	578	24	AAS17900	DNA encoding bovin
	4	492.4	83.2	556	24	AAS17901	Partail cDNA encod
	5	476.8	80.5	1023	24	ABN96895	Gene #3393 used to
	6	476.8	80.5	1023	24	ABL61914	Colon adenocarcino
	7	476.8	80.5	1023	24	ABL64458	Stomach cancer rel
	8	476.8	80.5	1023	24	ABL64961	Lung cancer relate
	9	476.8	80.5	1079	18	AAT74076	Human LS-DNase nuc
	10	475.2	80.3	1108	24	ABK84429	Human cDNA differe
	11	451.2	76.2	1208	17	AAT17733	Deoxyribonuclease,
	12	435.2	73.5	1124	18	AAT74083	Murine LS-DNase nu
	13	348.4	58.9	350	24	AAS17902	partial cDNA encod
	14	314.8	53.2	350	24	AAS17903	partial cDNA encod
	15	300.6	50.8	349	24	AAS17904	Partial cDNA encod
c	16	171.4	29.0	392	24	ABN94132	Gene #630 used to
c	17	171.4	29.0	392	24	ABL62246	Colon adenocarcino
	18	149.4	25.2	1039	15	AAQ54429	Human DNase I. Ho
	19	149.4	25.2	1039	22	ABA02661	Human DNase I 1.
	20	147.8	25.0	783	22	ABA02662	Human DNase I 2.
	21	147.8	25.0	858	22	ABA02663	Human DNase I 3.
	22	147.8	25.0	1000	21	AAA28433	DNA encoding STII
	23	147.8	25.0	1039	11	AAQ05512	Sequence encoding
	24	147.8	25.0	1039	15	AAQ62645	Human DNase gene.
	25	147.8	25.0	1548	22	ABA02702	Humanised HMFG-1 F
	26	147.8	25.0	1548	22	ABA02703	Humanised HMFG-1 F
	27	147.8	25.0	1548	22	ABA02705	Humanised HMFG-1 F
	28	147.8	25.0	1554	22	ABA02682	Humanised HMFG-1 F
	29	147.8	25.0	1554	22	ABA02683	Humanised HMFG-1 F
	30	147.8	25.0	1554	22	ABA02685	Humanised HMFG-1 F
	31	147.8	25.0	1557	22	ABA02704	Humanised HMFG-1 F
	32	147.8	25.0	1560	22	ABA02710	Humanised HMFG-1 F
	33	147.8	25.0	1560	22	ABA02711	Humanised HMFG-1 F
	34	147.8	25.0	1560	22	ABA02713	Humanised HMFG-1 F
	35	147.8	25.0	1560	22	ABA02714	Humanised HMFG-1 F
	36	147.8	25.0	1560	22	ABA02715	Humanised HMFG-1 F
	37	147.8	25.0	1560	22	ABA02716	Humanised HMFG-1 F
	38	147.8	25.0	1563	22	ABA02684	Humanised HMFG-1 F
	39	147.8	25.0	1566	22	ABA02706	Humanised HMFG-1 F
	40	147.8	25.0	1566	22	ABA02707	Humanised HMFG-1 F
	41	147.8	25.0	1566	22	ABA02709	Humanised HMFG-1 F
	42	147.8	25.0	1569	22	ABA02712	Humanised HMFG-1 F
	43	147.8	25.0	1575	22	ABA02708	Humanised HMFG-1 F
	44	147.8	25.0	1578	22	ABA02717	Humanised HMFG-1 F
	45	147.8	25.0	1578	22	ABA02718	Humanised HMFG-1 F

ALIGNMENTS

RESULT 1
AAD29088
ID AAD29088 standard; cDNA; 592 BP.
XX
AC AAD29088;
XX
DT 16-MAY-2002 (first entry)
XX
DE Bovine fertility associated antigen (FAA) partial cDNA.
XX
KW Bovine; fertility associated antigen; FAA; sperm fertility; acrosome;
KW artificial insemination; plasma membrane; sperm cell; anti-infertility;
KW reproductive tract; ss.
XX
OS Bos sp.
XX
FH Key Location/Qualifiers
FT CDS 1..591
FT /*tag= a
FT /product= "Bovine fertility associated antigen (FAA)"
FT /note= "CDS does not include start and stop codon"
FT /partial
XX
PN WO200206444-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US22097.
XX
PR 14-JUL-2000; 2000US-218140P.
XX
PA (ARIZ-) ARIZONA BOARD OF REGENTS.

XX
PI Zhang H, Ax RL, Bellin ME;
XX
DR WPI; 2002-171803/22.
DR P-PSDB; AAE18281.
XX
PT Novel amino acid and polynucleotide sequence for fertility associated
PT antigen useful for increasing stability of plasma membrane, acrosome
PT and other portions of sperm cell, and increasing fertility of mammals
PT -
XX
PS Claim 1; Fig 6; 37pp; English.
XX
CC The invention relates to bovine fertility associated antigen (FAA) and
CC its nucleic acid sequence. FAA is useful for increasing the stability of
CC the plasma membrane plus acrosome of a sperm cell and/or other portions
CC of a sperm cell, in particular mammalian sperm cell such as a sperm cell
CC from buffalo, cow, horse, mice, pig, sheep, human, avian sperm cell
CC including turkey or chicken sperm cell. FAA is also useful for increasing
CC the fertility of sperm from a mammal including buffalo, cow, horse, mice,
CC pig, sheep or human, or avian (chicken or turkey). FAA is added to the
CC suspension of sperm cells prior to preservation or prior to
CC administration of the sperm cells to a mammal in artificial insemination.
CC Alternatively, fertility of sperm from a mammal is increased by placing
CC FAA into a female's reproductive tract prior to deposition of the sperm
CC cells into the female's reproductive tract by copulation or artificial
CC insemination. The fertility of the male mammal is increased by injecting
CC FAA into the male mammal's reproductive tract. The present sequence is
CC bovine fertility associated antigen (FAA) partial cDNA.
XX
SQ Sequence 592 BP; 170 A; 143 C; 142 G; 137 T; 0 other;

Query Match 100.0%; Score 592; DB 24; Length 592;
Best Local Similarity 100.0%; Pred. No. 6e-178;
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
```

RESULT 2
AAD29089
ID AAD29089 standard; DNA; 671 BP.
XX
AC AAD29089;
XX
DT 16-MAY-2002 (first entry)
XX
DE Bovine 22kDa recombinant FAA (rFAA) coding sequence.
XX
KW Bovine; fertility associated antigen; FAA; sperm fertility; acrosome;
KW artificial insemination; plasma membrane; sperm cell; anti-infertility;
KW reproductive tract; ds.
XX
OS Bos sp.
XX
PN WO200206444-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US22097.
XX
PR 14-JUL-2000; 2000US-218140P.
XX
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
PI Zhang H, Ax RL, Bellin ME;
XX
DR WPI; 2002-171803/22.
XX
PT Novel amino acid and polynucleotide sequence for fertility associated
PT antigen useful for increasing stability of plasma membrane, acrosome
PT and other portions of sperm cell, and increasing fertility of mammals
PT -
XX
PS Claim 1; Fig 5; 37pp; English.
XX
CC The invention relates to bovine fertility associated antigen (FAA) and
CC its nucleic acid sequence. FAA is useful for increasing the stability of
CC the plasma membrane plus acrosome of a sperm cell and/or other portions
CC of a sperm cell, in particular mammalian sperm cell such as a sperm cell
CC from buffalo, cow, horse, mice, pig, sheep, human, avian sperm cell
CC including turkey or chicken sperm cell. FAA is also useful for increasing
CC the fertility of sperm from a mammal including buffalo, cow, horse, mice,
CC pig, sheep or human, or avian (chicken or turkey). FAA is added to the
CC suspension of sperm cells prior to preservation or prior to
CC administration of the sperm cells to a mammal in artificial insemination.
CC Alternatively, fertility of sperm from a mammal is increased by placing
CC FAA into a female's reproductive tract prior to deposition of the sperm
CC cells into the female's reproductive tract by copulation or artificial
CC insemination. The fertility of the male mammal is increased by injecting
CC FAA into the male mammal's reproductive tract. The present sequence is
CC bovine 22kDa recombinant FAA (rFAA) coding sequence.
XX
SQ Sequence 671 BP; 195 A; 166 C; 154 G; 156 T; 0 other;

Query Match 100.0%; Score 592; DB 24; Length 671;
Best Local Similarity 100.0%; Pred. No. 6.5e-178;
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     27 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 86

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 146

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    147 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 206

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 266

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 326

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 386

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 446

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 506

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    507 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 566

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    567 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 618
```

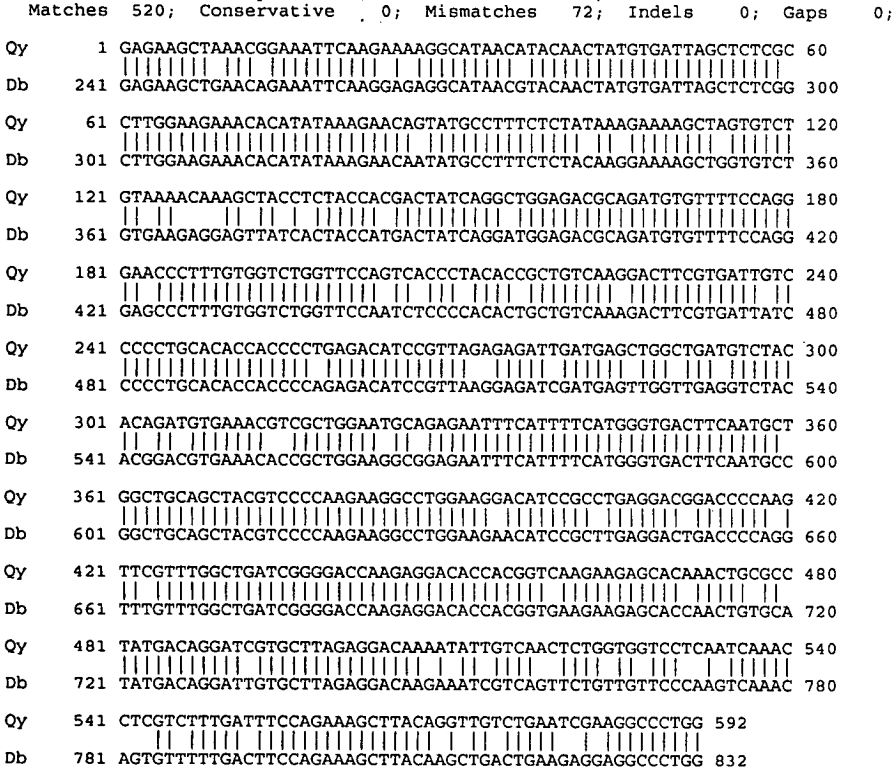
```
RESULT 3
AAS17900
ID   AAS17900 standard; cDNA; 578 BP.
XX
AC   AAS17900;
XX
DT   26-MAR-2002  (first entry)
XX
DE   DNA encoding bovine fertility-associated antigen (FAA).
XX
KW   Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;
KW   reproductive fitness; single nucleotide polymorphism; SNP; bovine; cow;
KW   ss.
XX
OS   Bos taurus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..578
FT                   /*tag=  a
FT                   /product= "FAA"
FT                   /note= "Fertility-associated antigen"
FT                   /partial
FT                   /note= "No start or stop codon given"
FT                   /transl_except= (pos:289..300, aa:A)
FT                   /note= "This codon has an apparent 9 nucleotide
FT                         insertion"
FT                   /transl_except= (pos:325..326, aa:R)
FT                   /note= "This codon has an apparent 1 nucleotide
FT                          deletion which alters the reading frame"
FT                   /transl_except= (pos:576..578, aa:RLSESKAL)
XX
PN   WO200176529-A2.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-US10802.
XX
PR   07-APR-2000; 2000US-195225P.
XX
PA   (ZHAN/) ZHANG H M.
PA   (AXRL/) AX R L.
XX
PI   Zhang HM,  Ax RL;
XX
DR   WPI; 2002-066307/09.
DR   P-PSDB; AAU11516.
XX
PT   Novel human or bovine fertility-associated antigen useful for
PT   stabilising sperm cell acrosome and increasing fertility of a male
XX
PS   Claim 25; Fig 2A-E; 54pp; English.
XX
CC   The invention describes a novel isolated human fertility-associated
CC   antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC   for increasing the stability of a sperm cell acrosome and for increasing
CC   the fertility of a human male by administering FAA. Detecting single
CC   nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC   for assaying fertility by detecting the presence or absence of a FAA SNP
CC   in the mammal and correlating the presence or absence with the fertility
CC   of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC   specification. FAA SNP is also useful for determining the reproductive
CC   fitness of a human or bovine, by determining the nucleotide sequence of
CC   the FAA SNP, quantifying the fertility of more than one mammal containing
CC   the FAA SNP and correlating the frequency of the FAA SNP to the
CC   reproductive fitness of the mammals. FAA improves the integrity of sperm
CC   membranes and increases the capacitation of sperm derived from either
CC   fertile or infertile humans. Early genetic identification of infertility
CC   improves the prognosis for subsequent attempts at fertilisation and
CC   facilitates early intervention to determine whether the individual may
CC   benefit from fertility treatments to avoid costly and/or emotional
CC   problems with attempted inseminations. This sequence encodes the bovine
CC   fertility-associated antigen (FAA) described in the method of the
CC   invention.
XX
SQ   Sequence 578 BP; 167 A; 140 C; 138 G; 133 T; 0 other;
```

```
  Query Match             91.0%;  Score 539;  DB 24;  Length 578;
  Best Local Similarity   98.3%;  Pred. No. 4.4e-161;
  Matches  569;  Conservative    0;  Mismatches    0;  Indels   10;  Gaps    2;

Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGA---------TGAGCTGGCT 291
          |||||||||||||||||||||||||||||||||||||||||         ||||||||||
Db    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGAGCTGGCT 300

Qy    292 GATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGAC 351
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    301 GATGTCTACACAGATGTGAAACGTC-CTGGAATGCAGAGAATTTCATTTTCATGGGTGAC 359

Qy    352 TTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACG 411
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 TTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACG 419

Qy    412 GACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACA 471
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 GACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACA 479

Qy    472 AACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCT 531
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    480 AACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCT 539

Qy    532 CAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570
          ||||||||||||||||||||||||||||||||||||||
Db    540 CAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 578


RESULT 4
AAS17901
ID   AAS17901 standard; cDNA; 556 BP.
XX
AC   AAS17901;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Partail cDNA encoding human fertility-associated antigen (FAA), HC1.
XX
KW   Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;
KW   reproductive fitness; single nucleotide polymorphism; SNP; human;
KW   HC1; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..555
FT                   /*tag=  a
FT                   /product= "Human FAA"
FT                   /note= "Human fertility-associated antigen, HC1"
FT                   /partial
FT                   /note= "No start or stop codon given"
FT                   /transl_except= (pos:1..3, aa:KVIKRCDIILVMEIKDS)
FT                   /transl_except= (pos: 253..255, aa:S)
XX
PN   WO200176529-A2.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-US10802.
XX
PR   07-APR-2000; 2000US-195225P.
XX
PA   (ZHAN/) ZHANG H M.
PA   (AXRL/) AX R L.
XX
PI   Zhang HM,  Ax RL;
XX
DR   WPI; 2002-066307/09.
DR   P-PSDB; AAU11517.
XX
PT   Novel human or bovine fertility-associated antigen useful for
PT   stabilising sperm cell acrosome and increasing fertility of a male  -
XX
PS   Claim 1; Fig 2A-E; 54pp; English.
XX
CC   The invention describes a novel isolated human fertility-associated
CC   antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC   for increasing the stability of a sperm cell acrosome and for increasing
CC   the fertility of a human male by administering FAA. Detecting single
CC   nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC   for assaying fertility by detecting the presence or absence of a FAA SNP
CC   in the mammal and correlating the presence or absence with the fertility
CC   of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC   specification. FAA SNP is also useful for determining the reproductive
CC   fitness of a human or bovine, by determining the nucleotide sequence of
CC   the FAA SNP, quantifying the fertility of more than one mammal containing
CC   the FAA SNP and correlating the frequency of the FAA SNP to the
CC   reproductive fitness of the mammals. FAA improves the integrity of sperm
CC   membranes and increases the capacitation of sperm derived from either
CC   fertile or infertile humans. Early genetic identification of infertility
CC   improves the prognosis for subsequent attempts at fertilisation and
CC   facilitates early intervention to determine whether the individual may
CC   benefit from fertility treatments to avoid costly and/or emotional
CC   problems with attempted inseminations. This sequence encodes the human
CC   fertility-associated antigen (FAA), HC1, one of 4 partial prostate cDNA
CC   clones of the novel FAA gene described in the method of the invention.
XX
SQ   Sequence 556 BP; 159 A; 130 C; 142 G; 125 T; 0 other;

  Query Match           83.2%;  Score 492.4;  DB 24;  Length 556;
  Best Local Similarity 96.0%;  Pred. No. 3e-146;
  Matches  505;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||
Db     31 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 90

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||| ||||||||||||| || ||||||| ||||||
Db     91 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 150

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db    151 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 210

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    211 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 270

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    271 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 330

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    331 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 390

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    391 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 450

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 510

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTG 526
          |||||||||||||||||||||||||||||||||||||||||||| ||
Db    511 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGTTG 556


RESULT 5
ABN96895
ID   ABN96895 standard; DNA; 1023 BP.
XX
AC   ABN96895;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #3393 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US30589.
XX
PR   02-OCT-2000; 2000US-237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
```

XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 3393; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

Query Match 80.5%; Score 476.8; DB 24; Length 1023;
Best Local Similarity 87.8%; Pred. No. 3.8e-141;
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;



RESULT 6
ABL61914
ID ABL61914 standard; DNA; 1023 BP.
XX
AC ABL61914;
XX
DT 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:251.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX

PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 251; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

Query Match 80.5%; Score 476.8; DB 24; Length 1023;
Best Local Similarity 87.8%; Pred. No. 3.8e-141;
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||
Db    241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db    301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db    361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db    421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          |||||||||||||||| ||||||||||||  ||||| |||||| ||| ||| ||||||
Db    481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db    541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db    601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || ||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db    661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||| | || ||||  |||| || |||  | ||||||
Db    721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            || ||||| |||||||||||||||| | || |||||  | |||||||||
Db    781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832
```

RESULT 7
ABL64458

ID ABL64458 standard; DNA; 1023 BP.
XX
AC ABL64458;
XX
DT 15-MAY-2002 (first entry)
XX
DE Stomach cancer related gene sequence SEQ ID NO:2795.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and

```
PT   determining a change in expression of a gene of a signature gene set  -
XX
PS   Claim 1; SEQ ID 2795; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an
CC   anti-neoplastic agent. The method involves exposing cells to a chemical
CC   agent to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening
CC   an anti-neoplastic agent, and can be used for producing a product which
CC   is the data collected with respect to the anti-neoplastic agent as a
CC   result of M1, and the data is sufficient to convey the chemical
CC   structure and/or properties of the agent. M1 can be used in the
CC   treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC   oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC   adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC   infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC   carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ   Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

  Query Match             80.5%;  Score 476.8;  DB 24;  Length 1023;
  Best Local Similarity   87.8%;  Pred. No. 3.8e-141;
  Matches  520;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
Db     241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
Db     301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
Db     361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
Db     421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
Db     481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
Db     541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
Db     601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
Db     661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
Db     721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
Db     781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832


RESULT 8
ABL64961
ID   ABL64961 standard; DNA; 1023 BP.
XX
AC   ABL64961;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Lung cancer related gene sequence SEQ ID NO:3298.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US10838.
XX
PR   05-JUN-2000; 2000US-209473P.
PR   05-JUN-2000; 2000US-209531P.
PR   18-SEP-2000; 2000US-233133P.
PR   18-SEP-2000; 2000US-233617P.
PR   20-SEP-2000; 2000US-234009P.
PR   20-SEP-2000; 2000US-234034P.
PR   20-SEP-2000; 2000US-234052P.
PR   22-SEP-2000; 2000US-234509P.
PR   22-SEP-2000; 2000US-234567P.
PR   25-SEP-2000; 2000US-234923P.
PR   25-SEP-2000; 2000US-234924P.
PR   25-SEP-2000; 2000US-235077P.
PR   25-SEP-2000; 2000US-235082P.
PR   25-SEP-2000; 2000US-235134P.
PR   25-SEP-2000; 2000US-235280P.
PR   26-SEP-2000; 2000US-235637P.
PR   26-SEP-2000; 2000US-235638P.
PR   27-SEP-2000; 2000US-235711P.
PR   27-SEP-2000; 2000US-235720P.
PR   27-SEP-2000; 2000US-235840P.
PR   27-SEP-2000; 2000US-235863P.
PR   28-SEP-2000; 2000US-236028P.
PR   28-SEP-2000; 2000US-236032P.
PR   28-SEP-2000; 2000US-236033P.
PR   28-SEP-2000; 2000US-236034P.
PR   28-SEP-2000; 2000US-236109P.
PR   28-SEP-2000; 2000US-236111P.
PR   29-SEP-2000; 2000US-236842P.
PR   29-SEP-2000; 2000US-236891P.
PR   02-OCT-2000; 2000US-237172P.
PR   02-OCT-2000; 2000US-237173P.
PR   02-OCT-2000; 2000US-237278P.
PR   02-OCT-2000; 2000US-237294P.
PR   02-OCT-2000; 2000US-237295P.
PR   02-OCT-2000; 2000US-237316P.
PR   03-OCT-2000; 2000US-237425P.
PR   03-OCT-2000; 2000US-237598P.
PR   03-OCT-2000; 2000US-237604P.
PR   03-OCT-2000; 2000US-237606P.
PR   03-OCT-2000; 2000US-237608P.
PR   01-NOV-2000; 2000US-244867P.
PR   01-NOV-2000; 2000US-245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE,  Augustus M,  Carter KC,  Ebner R,  Endress G,  Horrigan S;
PI   Soppet DR,  Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a
PT   chemical agent to be tested for anti-neoplastic activity, and
PT   determining a change in expression of a gene of a signature gene set  -
XX
```

```
PS   Claim 1; SEQ ID 3298; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an
CC   anti-neoplastic agent. The method involves exposing cells to a chemical
CC   agent to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening
CC   an anti-neoplastic agent, and can be used for producing a product which
CC   is the data collected with respect to the anti-neoplastic agent as a
CC   result of M1, and the data is sufficient to convey the chemical
CC   structure and/or properties of the agent. M1 can be used in the
CC   treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC   oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC   adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC   infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC   carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ   Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

  Query Match             80.5%;  Score 476.8;  DB 24;  Length 1023;
  Best Local Similarity   87.8%;  Pred. No. 3.8e-141;
  Matches  520;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
           |||||||| ||| |||||||||| | |||||||| ||||||||||||||||||||||||
Db     241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
           |||||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db     301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
           || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db     361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
           || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
           |||||||||||||||| ||||||||||||  ||||| |||||| ||| ||| ||||||
Db     481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
           || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||||
Db     541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
           |||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db     601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
           || |||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db     661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
           |||||||||| |||||||||||||||| | || ||||  |||| || |||  | ||||||
Db     721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
             || ||||| |||||||||||||||| | || |||||  | |||||||||
Db     781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832


RESULT 9
AAT74076
ID   AAT74076 standard; cDNA; 1079 BP.
XX
AC   AAT74076;
XX
DT   29-JAN-1998  (first entry)
XX
DE   Human LS-DNase nucleotide sequence.
XX
KW   DNase; actin; DNA viscoelasticity; systemic lupus erythematosus;
KW   cystic fibrosis; meningitis; pulmonary disease; gene therapy;
KW   recombinant preparation; DNA-hydrolytic activity; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             71..988
FT                   /*tag=  a
FT   sig_peptide     71..130
FT                   /*tag=  b
FT   mat_peptide     131..985
FT                   /*tag=  c
FT                   /product= LS-DNase
XX
PN   WO9728266-A1.
XX
PD   07-AUG-1997.
XX
PF   03-FEB-1997;   97WO-US01506.
XX
PR   05-FEB-1996;   96US-0597078.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Baron WF;
XX
DR   WPI; 1997-402619/37.
DR   P-PSDB; AAW23542.
XX
PT   New isolated LS-DNase which is resistant to inhibition by actin -
PT   used for the treatment of e.g systemic lupus erythematosus and
PT   pulmonary diseases and disorders such as cystic fibrosis
XX
PS   Claim 1; Fig 1; 32pp; English.
XX
CC   This DNA sequence encodes a novel human LS-DNase (AAW23542), which has
CC   DNA-hydrolytic activity but is resistant to inhibition by actin. It can
CC   be used for reducing the viscoelasticity or viscous consistency of
CC   DNA-containing material, as well as treatment of systemic lupus
CC   erythematosus or pulmonary disease of disorders such as cystic fibrosis,
CC   bronchitis, pneumonia, bronchiectasis, emphysema, asthma, tuberculosis
CC   or fungal infections. It can also be used for adjunctive treatment of
CC   abscesses or severe closed-space infections in conditions such as
CC   empyema, meningitis, peritonitis, sinusitis, otitis, periodontitis,
CC   pericarditis, pancreatitis, cholelithiasis, endocarditis and
CC   septic arthritis, as well as in topical treatments in a variety of
CC   inflammatory and infected lesions, e.g. of the skin and mucosal
CC   membranes, surgical wounds, ulcerative lesions and burns. LS-DNase may
CC   improve the efficacy of antibiotics used in the treatment of such
CC   infections (e.g. gentamicin activity is markedly reduced by reversible
CC   binding to intact DNA). It can also be used for the treatment of other
CC   non-infected conditions in which there is an accumulation of cellular
CC   debris that includes cellular DNA, such as pyelonephritis and
CC   tubulo-interstitial kidney disease. It can also be used in in vitro
CC   diagnostic assays of a viscous material (e.g. sputum) from a patient.
CC   The anti-LS-DNase antibodies can be used for detection and purification
CC   of LS-DNase.
XX
SQ   Sequence 1079 BP; 330 A; 260 C; 252 G; 237 T; 0 other;

  Query Match             80.5%;  Score 476.8;  DB 18;  Length 1079;
  Best Local Similarity   87.8%;  Pred. No. 3.9e-141;
  Matches  520;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
           |||||||| ||| |||||||||| | |||||||| ||||||||||||||||||||||||
Db     287 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 346
```

```
Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
           ||||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db     347 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 406

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
           || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db     407 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 466

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
           || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     467 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 526

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
           ||||||||||||||||| ||||||||||||  ||||| |||||| ||| ||| ||||||
Db     527 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 586

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
           || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db     587 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 646

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
           |||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db     647 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 706

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
           || ||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db     707 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 766

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
           |||||||||| |||||||||||||| | || |||| |||| || |||  | ||||||
Db     767 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 826

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
              || ||||| |||||||||||||||| | || |||||  | |||||||||
Db     827 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 878
```

RESULT 10
ABK84429
ID ABK84429 standard; cDNA; 1108 BP.
XX
AC ABK84429;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #1000.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US30821.
XX
PR 03-OCT-2000; 2000US-237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity -
XX
PS Claim 1; SEQ ID No 1000; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1108 BP; 329 A; 261 C; 271 G; 247 T; 0 other;

Query Match 80.3%; Score 475.2; DB 24; Length 1108;
Best Local Similarity 87.7%; Pred. No. 1.3e-140;
Matches 519; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

```
Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
           |||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||||
Db     309 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 368

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
           |||||||||| |||||||||||||| |||||||||||||| || |||||||| ||||||
Db     369 CTTGGAAGAAAAACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 428

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
           || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db     429 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 488

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
           || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     489 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 548

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
           ||||||||||||||||| |||||||||||||  ||||| |||||| ||| ||| ||||||
Db     549 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 608

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
           || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db     609 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 668
```

```
Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          ||||||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db    669 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 728

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || ||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db    729 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 788

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||| | || |||| |||| || ||| | ||||||
Db    789 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 848

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            || ||||| |||||||||||||||| | || |||||  | |||||||||
Db    849 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 900
```

RESULT 11
AAT17733
ID AAT17733 standard; cDNA; 1208 BP.
XX
AC AAT17733;
XX
DT 17-OCT-1996 (first entry)
XX
DE Deoxyribonuclease, DNase-gamma cDNA.
XX
KW Deoxyribonuclease; DNase-gammma; selective cleavage; linker;
KW chromatin DNA; cell nucleus; 3'-hydroxy; 5'-phosphate; apoptosis;
KW carcinogenesis; AIDS development; diagnosis; prevention;
KW treatment; cancer; autoimmune disease; viral infections; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 12..944
FT /*tag= a
FT /note= "GAT codon comprising bases 837-839 has
FT no corresponding amino acid in AAR94021"
XX
PN WO9607735-A1.
XX
PD 14-MAR-1996.
XX
PF 06-SEP-1995; 95WO-JP01775.
XX
PR 06-SEP-1994; 94JP-0239518.
XX
PA (TANU/) TANUMA S.
XX
PI Tanuma S;
XX
DR WPI; 1996-171610/17.
DR P-PSDB; AAR94021.
XX
PT DNase which selectively cuts linker region of chromatin DNA -
PT useful in diagnosis, treatment and prevention of cancer, autoimmune
PT diseases, viral infections, etc.
XX
PS Claim 4; Pages 55-57; 75pp; Japanese.
XX
CC The present sequence encodes the deoxyribonuclease, DNase-gammma,
CC which is able to selectively cleave the linker part of chromatin
CC DNA. DNase-gamma has a mol. wt. of 32000 by SDS-PAGE, or 30000 by
CC gel filtration, an optimum pH of 5.6, is found in the cell nucleus,
CC is not dependent on divalent cations, has a Zn(2+) inhibition
CC IC(50) > 1mM and leaves 3'-OH and 5'-P termini after cleavage. The
CC DNase can be used at a mol. level to resolve the control
CC mechanisms of carcinogenesis, the autoimmune system and AIDS
CC development. It can also be used in the diagnosis, prevention and
CC treatment of cancer, autoimmune disease, viral infections, etc.,
CC and as an apoptosis control agent.
XX

SQ Sequence 1208 BP; 319 A; 304 C; 282 G; 303 T; 0 other;

Query Match 76.2%; Score 451.2; DB 17; Length 1208;
Best Local Similarity 85.1%; Pred. No. 5.8e-133;
Matches 504; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```
Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| || ||||| ||| ||| | |||  |||||||||| ||||||||||||||
Db    243 GAGAAGCTGAATGGAAACTCACGAAGAAGCACGACATACAACTACGTGATTAGCTCTCGG 302

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||||||||||||| ||||| || || ||||| ||||||
Db    303 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTCCTCTACAAGGAGAAGCTGGTGTCT 362

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||   || |||||||||| |||||||||| ||||||| |||| ||||||||||||
Db    363 GTGAAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 422

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||| |||||||| | |||| ||| |||| ||||||||||||||||||||
Db    423 GAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCACTGCTGCCAAGGACTTCGTGATTGTC 482

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||| ||||||| || ||||| || || |||| |||||| |||||||||||||| ||||||
Db    483 CCCTTGCACACAACTCCTGAAACCTCTGTTAAAGAGATAGATGAGCTGGCTGACGTCTAC 542

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || ||||| | | | || ||||| ||||||| |||||| |||||||||| |||||||||
Db    543 ACGGATGTTAGAAGACGATGGAAGGCAGAGATTTTCATCTTCATGGGTGATTTCAATGCT 602

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          ||||||||||||||||||||||||||||||||||| |||||||  ||||||| ||||||||
Db    603 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGTTTGAGGACAGACCCCAAC 662

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||| ||||||||||||||||||||||||||||||||||||||| | ||| |||
Db    663 TTTGTTTGGCTGATTGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACCAGCTGTGCC 722

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| ||  || | |||||| | || |||||||||  | ||| |  || |
Db    723 TATGACAGGATTGTCGTTCGCGGACAAGAGATAGTCAACTCTGTGGTTCCCCGCTCCAGT 782

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            |||||||| || ||||||||||   |||||||||  | ||||||||
Db    783 GGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGG 834
```

RESULT 12
AAT74083
ID AAT74083 standard; cDNA; 1124 BP.
XX
AC AAT74083;
XX
DT 29-JAN-1998 (first entry)
XX
DE Murine LS-DNase nucleotide sequence.
XX
KW DNase; actin; DNA viscoelasticity; systemic lupus erythematosus;
KW cystic fibrosis; meningitis; pulmonary disease; gene therapy;
KW recombinant preparation; DNA-hydrolytic activity; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 173..1105
FT /*tag= a
FT sig_peptide 173..247
FT /*tag= b
FT /note= "putative"
FT mat_peptide 248..1105
FT /*tag= c
FT /product= LS-DNase
XX

```
PN   WO9728266-A1.
XX
PD   07-AUG-1997.
XX
PF   03-FEB-1997;   97WO-US01506.
XX
PR   05-FEB-1996;   96US-0597078.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Baron WF;
XX
DR   WPI; 1997-402619/37.
XX
PT   New isolated LS-DNase which is resistant to inhibition by actin -
PT   used for the treatment of e.g systemic lupus erythematosus and
PT   pulmonary diseases and disorders such as cystic fibrosis
XX
PS   Example 6; Fig 3; 32pp; English.
XX
CC   This DNA sequence encodes a variant LS-DNase, which has DNA-hydrolytic
CC   activity but is resistant to inhibition by actin. A 649 base pair
CC   fragment of the coding sequence of cloned human LS-DNase (AAT74076) was
CC   used to screen a murine liver cDNA library and obtain this nucleotide.
CC   Murine mature LS-DNase is 84% identical to human LS-DNase. A portion of
CC   the present sequence was used in Northern blot assays of various murine
CC   tissues. Expression of murine LS-DNase mRNA was found to be highest in
CC   the liver and spleen. Human LS-DNase can be used for reducing the
CC   viscoelasticity or viscous consistency of DNA-containing material, as
CC   well as treatment of systemic lupus erythematosus or pulmonary disease
CC   disorders such as cystic fibrosis. It can also be used for adjunctive
CC   treatment of abscesses or severe closed-space infections, e.g.
CC   meningitis. It may improve the efficacy of antibiotics used in the
CC   treatment of such infections (e.g. gentamicin activity is markedly
CC   reduced by reversible binding to intact DNA). It can also be used in in
CC   vitro diagnostic assays of a viscous material (e.g. sputum) from a
CC   patient. The anti-LS-DNase antibodies can be used for detection and
CC   purification of LS-DNase.
XX
SQ   Sequence 1124 BP; 299 A; 273 C; 273 G; 279 T; 0 other;

  Query Match             73.5%;  Score 435.2;  DB 18;  Length 1124;
  Best Local Similarity   83.4%;  Pred. No. 6.9e-128;
  Matches  494;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
Db     404 GAGAAGCTGAATGGAAATTCACGAAGAAGCACAACATACAACTATGTGATTAGTTCTCGA 463

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
Db     464 CTTGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGCTGGTGTCT 523

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
Db     524 GTGAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 583

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
Db     584 GAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCGTGATTGTC 643

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
Db     644 CCCTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCGATGTCTAC 703

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
Db     704 ACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATTTCAACGCC 763

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
Db     764 GGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGAGGACGGACCCCAAG 823

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
Db     824 TTTGTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCAGCTGTGCC 883

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
Db     884 TATGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTTCCCCGTTCCAGT 943

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
Db     944 GGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 995


RESULT 13
AAS17902
ID   AAS17902 standard; cDNA; 350 BP.
XX
AC   AAS17902;
XX
DT   26-MAR-2002  (first entry)
XX
DE   partial cDNA encoding human fertility-associated antigen (FAA), HC2.
XX
KW   Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;
KW   reproductive fitness; single nucleotide polymorphism; SNP; human;
KW   HC2; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..349
FT                   /*tag=  a
FT                   /product= "Human FAA"
FT                   /note= "Human fertility-associated antigen, HC2"
FT                   /partial
FT                   /note= "No start or stop codon given"
XX
PN   WO200176529-A2.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-US10802.
XX
PR   07-APR-2000; 2000US-195225P.
XX
PA   (ZHAN/) ZHANG H M.
PA   (AXRL/) AX R L.
XX
PI   Zhang HM,  Ax RL;
XX
DR   WPI; 2002-066307/09.
DR   P-PSDB; AAU11518.
XX
PT   Novel human or bovine fertility-associated antigen useful for
PT   stabilising sperm cell acrosome and increasing fertility of a male  -
XX
PS   Claim 1; Fig 2A-E; 54pp; English.
XX
CC   The invention describes a novel isolated human fertility-associated
CC   antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC   for increasing the stability of a sperm cell acrosome and for increasing
CC   the fertility of a human male by administering FAA. Detecting single
CC   nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC   for assaying fertility by detecting the presence or absence of a FAA SNP
CC   in the mammal and correlating the presence or absence with the fertility
CC   of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC   specification. FAA SNP is also useful for determining the reproductive
CC   fitness of a human or bovine, by determining the nucleotide sequence of
CC   the FAA SNP, quantifying the fertility of more than one mammal containing
CC   the FAA SNP and correlating the frequency of the FAA SNP to the
CC   reproductive fitness of the mammals. FAA improves the integrity of sperm
CC   membranes and increases the capacitation of sperm derived from either
CC   fertile or infertile humans. Early genetic identification of infertility
```

```
CC   improves the prognosis for subsequent attempts at fertilisation and
CC   facilitates early intervention to determine whether the individual may
CC   benefit from fertility treatments to avoid costly and/or emotional
CC   problems with attempted inseminations. This sequence encodes the human
CC   fertility-associated antigen (FAA), HC2, one of 4 partial prostate cDNA
CC   clones of the novel FAA gene described in the method of the invention.
XX
SQ   Sequence 350 BP; 88 A; 91 C; 93 G; 78 T; 0 other;

  Query Match             58.9%;  Score 348.4;  DB 24;  Length 350;
  Best Local Similarity   99.7%;  Pred. No. 1.6e-100;
  Matches  349;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    177 CAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGAT 236
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGAT 60

Qy    237 TGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGT 296
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGT 120

Qy    297 CTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAA 356
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAA 180

Qy    357 TGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCC 416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCC 240

Qy    417 CAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTG 476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTG 300

Qy    477 CGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTG 526
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    301 CGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGTTG 350


RESULT 14
AAS17903
ID   AAS17903 standard; cDNA; 350 BP.
XX
AC   AAS17903;
XX
DT   26-MAR-2002  (first entry)
XX
DE   partial cDNA encoding human fertility-associated antigen (FAA), HC3.
XX
KW   Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;
KW   reproductive fitness; single nucleotide polymorphism; SNP; human;
KW   HC3; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..349
FT                   /*tag=  a
FT                   /product= "Human FAA"
FT                   /note= "Human fertility associated antigen, HC3"
FT                   /partial
FT                   /note= "No start or stop codon given"
FT                   /transl_except= (pos:251..253, aa: Arg)
XX
PN   WO200176529-A2.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-US10802.
XX
PR   07-APR-2000; 2000US-195225P.
XX
PA   (ZHAN/) ZHANG H M.
PA   (AXRL/) AX R L.
XX
PI   Zhang HM,  Ax RL;
XX
DR   WPI; 2002-066307/09.
DR   P-PSDB; AAU11519.
XX
PT   Novel human or bovine fertility-associated antigen useful for
PT   stabilising sperm cell acrosome and increasing fertility of a male  -
XX
PS   Claim 1; Fig 2A-E; 54pp; English.
XX
CC   The invention describes a novel isolated human fertility-associated
CC   antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC   for increasing the stability of a sperm cell acrosome and for increasing
CC   the fertility of a human male by administering FAA. Detecting single
CC   nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC   for assaying fertility by detecting the presence or absence of a FAA SNP
CC   in the mammal and correlating the presence or absence with the fertility
CC   of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC   specification. FAA SNP is also useful for determining the reproductive
CC   fitness of a human or bovine, by determining the nucleotide sequence of
CC   the FAA SNP, quantifying the fertility of more than one mammal containing
CC   the FAA SNP and correlating the frequency of the FAA SNP to the
CC   reproductive fitness of the mammals. FAA improves the integrity of sperm
CC   membranes and increases the capacitation of sperm derived from either
CC   fertile or infertile humans. Early genetic identification of infertility
CC   improves the prognosis for subsequent attempts at fertilisation and
CC   facilitates early intervention to determine whether the individual may
CC   benefit from fertility treatments to avoid costly and/or emotional
CC   problems with attempted inseminations. This sequence encodes the human
CC   fertility-associated antigen (FAA), HC3, one of 4 partial prostate cDNA
CC   clones of the novel FAA gene described in the method of the invention.
XX
SQ   Sequence 350 BP; 89 A; 93 C; 94 G; 74 T; 0 other;

  Query Match             53.2%;  Score 314.8;  DB 24;  Length 350;
  Best Local Similarity   93.7%;  Pred. No. 7.9e-90;
  Matches  328;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy    177 CAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGAT 236
          |||||| |||||||||||||||||| || ||| |||| |||||||| ||||||||||
Db      1 CAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGAT 60

Qy    237 TGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGT 296
          | |||||||||||||||| |||||||||||| ||||| |||||| ||| ||| ||
Db     61 TATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGT 120

Qy    297 CTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAA 356
          ||||||  || |||||||  |||||||| || ||||||||||||||||||||||||||||
Db    121 CTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAA 180

Qy    357 TGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCC 416
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    181 TGCTGGCTGCAGCTACGTCCCCAAGAAGGCCCGGAAGGACATCCGCCTGAGGACGGACCC 240

Qy    417 CAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTG 476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTG 300

Qy    477 CGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTG 526
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    301 CGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGTTG 350


RESULT 15
AAS17904
ID   AAS17904 standard; cDNA; 349 BP.
XX
AC   AAS17904;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Partial cDNA encoding human fertility-associated antigen (FAA), HC4.
```

```
XX
KW   Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;
KW   reproductive fitness; single nucleotide polymorphism; SNP; human;
KW   HC4; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..348
FT                   /*tag= a
FT                   /product= "Human FAA"
FT                   /note= "Human fertility-associated antigen, HC4"
FT                   /partial
FT                   /note= "No start or stop codon given"
FT                   /transl_except= (pos: 11..12, aa:Phe)
FT                   /note= "This codon has an apparent 1 nucleotide
FT                          deletion which alters the reading frame"
XX
PN   WO200176529-A2.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-US10802.
XX
PR   07-APR-2000; 2000US-195225P.
XX
PA   (ZHAN/) ZHANG H M.
PA   (AXRL/) AX R L.
XX
PI   Zhang HM, Ax RL;
XX
DR   WPI; 2002-066307/09.
DR   P-PSDB; AAU11520.
XX
PT   Novel human or bovine fertility-associated antigen useful for
PT   stabilising sperm cell acrosome and increasing fertility of a male -
XX
PS   Claim 1; Fig 2A-E; 54pp; English.
XX
CC   The invention describes a novel isolated human fertility-associated
CC   antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC   for increasing the stability of a sperm cell acrosome and for increasing
CC   the fertility of a human male by administering FAA. Detecting single
CC   nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC   for assaying fertility by detecting the presence or absence of a FAA SNP
CC   in the mammal and correlating the presence or absence with the fertility
CC   of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC   specification. FAA SNP is also useful for determining the reproductive
CC   fitness of a human or bovine, by determining the nucleotide sequence of
CC   the FAA SNP, quantifying the fertility of more than one mammal containing
CC   the FAA SNP and correlating the frequency of the FAA SNP to the
CC   reproductive fitness of the mammals. FAA improves the integrity of sperm
CC   membranes and increases the capacitation of sperm derived from either
CC   fertile or infertile humans. Early genetic identification of infertility
CC   improves the prognosis for subsequent attempts at fertilisation and
CC   facilitates early intervention to determine whether the individual may
CC   benefit from fertility treatments to avoid costly and/or emotional
CC   problems with attempted inseminations. This sequence encodes the human
CC   fertility-associated antigen (FAA), HC4, one of 4 partial prostate cDNA
CC   clones of the novel FAA gene described in the method of the invention.
XX
SQ   Sequence 349 BP; 89 A; 91 C; 93 G; 76 T; 0 other;

  Query Match             50.8%;  Score 300.6;  DB 24;  Length 349;
  Best Local Similarity   91.6%;  Pred. No. 2.6e-85;
  Matches  318;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy    180 GGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGT 239
          || | || ||||||||||||||||| || ||| |||| |||||||| |||||||||||| |
Db      3 GGGAGCCCTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTAT 62

Qy    240 CCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTA 299
          ||||||||||||||||| ||||||||||||  ||||| |||||| ||| ||| |||||

Db     63 CCCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTA 122

Qy    300 CACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGC 359
          ||| || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db    123 CACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGC 182

Qy    360 TGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAA 419
          |||||||||||||||||||||||||||||||| |||||||| ||||||| |||||||
Db    183 CGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAG 242

Qy    420 GTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGC 479
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 GTTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGC 302

Qy    480 CTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTG 526
          |||||||||||||||||||||||||||||||||||||||||||||| ||
Db    303 CTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGTTG 349


Search completed: December  9, 2002, 09:17:55
Job time : 179.616 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:20:03 ; Search time 1320.4 Seconds
(without alignments)
7261.235 Million cell updates/sec

Title: US-09-905-114-1
Perfect score: 592
Sequence: 1 gagaagctaaacggaaattc.....gtctgaatcgaaggccctgg 592

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	478.4	80.8	890	9	AL546857	AL546857 AL546857
	2	476.8	80.5	896	9	AL546432	AL546432 AL546432
c	3	471.8	79.7	833	9	AL572037	AL572037 AL572037
c	4	469	79.2	931	9	AL571894	AL571894 AL571894
c	5	438.8	74.1	506	10	AV607154	AV607154 AV607154
	6	432	73.0	666	10	BE048177	BE048177 tz46h09.y
c	7	421	71.1	695	13	BI522352	BI522352 603081420
c	8	411.6	69.5	642	10	AW543356	AW543356 C0168H07-
	9	406	68.6	825	12	BG867772	BG867772 602786883
	10	396.4	67.0	825	12	BG530136	BG530136 602558626
c	11	391.2	66.1	620	10	AW543802	AW543802 C0174C07-
	12	389.4	65.8	628	13	BI460950	BI460950 603207187
c	13	386	65.2	610	10	AW540389	AW540389 C0132C04-
	14	383.8	64.8	668	10	AW915564	AW915564 EST346868
c	15	383.6	64.8	784	9	AI927844	AI927844 wo91g09.x
	16	376	63.5	832	13	BI838188	BI838188 603083104
c	17	363.4	61.4	634	9	AI660448	AI660448 we67c07.x
	18	363.4	61.4	981	14	BQ710632	BQ710632 AGENCOURT
	19	361	61.0	685	13	BI833952	BI833952 603085059
c	20	353.2	59.7	566	10	AW538998	AW538998 C0115F01-
	21	351	59.3	759	12	BG534290	BG534290 602553517
c	22	330.2	55.8	527	10	AW539589	AW539589 C0121F07-
c	23	328.2	55.4	613	10	AW779563	AW779563 hn81h08.x
	24	328	55.4	578	10	AW658627	AW658627 95214 MAR
c	25	311	52.5	643	10	AW001760	AW001760 ws04e04.x
	26	303.8	51.3	714	12	BG566893	BG566893 602589586
	27	302.2	51.0	855	12	BF237166	BF237166 602028119
c	28	296.8	50.1	542	12	BF434855	BF434855 7o74e07.x
	29	291.2	49.2	901	12	BF303508	BF303508 602030047
c	30	290	49.0	554	10	AW243180	AW243180 xn29e10.x
c	31	288.8	48.8	555	9	AI434430	AI434430 ti14c04.x
c	32	288.2	48.7	487	9	AI194886	AI194886 ui58a02.x
	33	279.8	47.3	341	10	AW803005	AW803005 IL2-UM007
c	34	279.6	47.2	497	9	AA564572	AA564572 nj25b03.s
	35	277.2	46.8	586	13	BM489658	BM489658 pgm2n.pk0
	36	267.6	45.2	522	14	BQ186387	BQ186387 UI-E-EJ1-
	37	265.6	44.9	595	12	BG077563	BG077563 H3016H03-
	38	262.6	44.4	453	14	H78170	H78170 yu84a11.r1
c	39	261.2	44.1	433	12	BF911384	BF911384 IL2-UT007
c	40	252.4	42.6	511	9	AI271665	AI271665 qj82a02.x
c	41	248.4	42.0	523	9	AI248904	AI248904 qh67e04.x
c	42	240.4	40.6	536	9	AA707043	AA707043 zj32d04.s
c	43	239.6	40.5	486	9	AA989198	AA989198 oq95d05.s
c	44	237.6	40.1	498	12	BG064211	BG064211 H3016H03-
	45	236.2	39.9	427	10	AW827162	AW827162 xn06e10.y

ALIGNMENTS

RESULT 1
AL546857
LOCUS AL546857 890 bp mRNA linear EST 16-FEB-2001
DEFINITION AL546857 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI026YK08 5 prime, mRNA sequence.
ACCESSION AL546857
VERSION AL546857.1 GI:12880381
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 890)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1..890
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI026YK08"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 250 a 215 c 220 g 204 t 1 others
ORIGIN

Query Match 80.8%; Score 478.4; DB 9; Length 890;
Best Local Similarity 88.0%; Pred. No. 1.1e-128;
Matches 521; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||||||||||||||||||||||||||
Db    216 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACATACAACTATGTGATTAGCTCTCGG 275

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db    276 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 335

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||||
Db    336 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 395

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db    396 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 455

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||| |||||||||||  ||||| |||||| ||| ||| ||||||
Db    456 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 515

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db    516 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 575

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||||| |||||||| ||||||| |||||||
Db    576 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 635

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db    636 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 695

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||| | || |||| |||| || ||| | ||||||
Db    696 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 755

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            || ||||| |||||||||||||| | || |||||  | |||||||||
Db    756 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 807
```

RESULT 2
AL546432
LOCUS AL546432 896 bp mRNA linear EST 16-FEB-2001
DEFINITION AL546432 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI030YD20 5 prime, mRNA sequence.
ACCESSION AL546432
VERSION AL546432.1 GI:12879540
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .896
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI030YD20"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 249 a 217 c 226 g 203 t 1 others
ORIGIN

Query Match 80.5%; Score 476.8; DB 9; Length 896;
Best Local Similarity 87.8%; Pred. No. 3.1e-128;
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | |||||||| |||||||||||||||||||||||||
Db    266 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 325

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db    326 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 385

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||||
Db    386 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 445

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db    446 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 505

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||| |||||||||||  ||||| |||||| ||| ||| ||||||
Db    506 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 565

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db    566 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 625

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||||| |||||||| ||||||| |||||||
Db    626 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 685

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db    686 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 745

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||| | || |||| |||| || ||| | ||||||
Db    746 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 805

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            || ||||| |||||||||||||| | || |||||  | |||||||||
Db    806 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 857
```

RESULT 3
AL572037/c
LOCUS AL572037 833 bp mRNA linear EST 16-FEB-2001
DEFINITION AL572037 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI026YK08 3 prime, mRNA sequence.

```
ACCESSION   AL572037
VERSION     AL572037.1  GI:12929919
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 833)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .833
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI026YK08"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      193 a    193 c    191 g    249 t      7 others
ORIGIN

  Query Match             79.7%;  Score 471.8;  DB 9;  Length 833;
  Best Local Similarity   87.0%;  Pred. No. 8.7e-127;
  Matches  515;  Conservative    2;  Mismatches   75;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
           |||||||| :|| |||||||||| | |||||| ||||||||||||||||||||||||||
Db     734 GAGAAGCTGWACAGAAATTCAAGGAGAGGCATYACATACAACTATGTGATTAGCTCTCGG 675

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
           ||||||||||:|||||||||||||| |||||||||||||| || |||||||| ||||||
Db     674 CTTGGAAGAAAMACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 615

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
           || ||    || || | |||||| |||||||||| |||| ||||||||||||||||||||
Db     614 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGANACGCAGATGTGTTTTCCAGG 555

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
           || |||||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     554 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 495

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
           |||||||||||||||| || |||||||||| ||||| |||||| ||| ||| ||||||
Db     494 CCCCTGCACACCACCCCAGAAACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 435

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
           || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db     434 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 375

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
           |||||||||||||||||||||||||||||||||||| |||||||| |||| || |||||||
Db     374 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGAACTGACCCCAGG 315

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
           || ||||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db     314 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 255

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
           |||||||||| |||||||||||||| | || |||| |||| || ||| | ||||||
Db     254 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 195

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
           || ||||| |||||||||||||||| | || ||||| | |||||||||
Db     194 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 143


RESULT 4
AL571894/c
LOCUS       AL571894                931 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL571894 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI030YD20 3
            prime, mRNA sequence.
ACCESSION   AL571894
VERSION     AL571894.1  GI:12929641
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 931)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .931
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI030YD20"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      209 a    219 c    225 g    271 t      7 others
ORIGIN

  Query Match             79.2%;  Score 469;  DB 9;  Length 931;
  Best Local Similarity   86.8%;  Pred. No. 6e-126;
  Matches  514;  Conservative    1;  Mismatches   77;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
           |||||||| ||| ||||||||||:| |||||||||| |||||||||||||| |||||||
Db     726 GAGAAGCTGAACAGAAATTCAAGRAGAGGCATAACGTACAACTATGTGATTGGCTCTCGG 667

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
           |||||||||||||||||||||||||| |||||||||||||| || |||| ||| ||||||
Db     666 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAGGCTGGTGTCT 607

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
           || ||    || ||   |||||| |||||||||| ||||||||||||||||||||||||
Db     606 GTGAAGAGGAGTTATNACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 547

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
           || |||||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     546 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 487

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
           |||||||||||||||| || |||||||||| ||||| |||||| ||| ||| ||||||
Db     486 CCCCTGCACACCACCCCAGAAACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 427

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
           || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
```

```
Db    426 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 367

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||||| |||||||| |||| || ||||||| |
Db    366 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGAACTGACCCCAGG 307

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db    306 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 247

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          ||| ||||||| |||||||||||||||| | || |||| |||| || ||| | |||||||
Db    246 TATAACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 187

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
             || ||||| |||||||||||||||| | || |||||  | |||||||||
Db    186 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 135
```

RESULT 5
AV607154/c

```
LOCUS       AV607154                506 bp    mRNA    linear   EST 28-NOV-2001
DEFINITION  AV607154 Bos taurus kidney fetus Bos taurus cDNA clone E1KI039B05
            3', mRNA sequence.
ACCESSION   AV607154
VERSION     AV607154.1  GI:9737527
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 506)
  AUTHORS   Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
            and Sugimoto,Y.
  TITLE     Establishment of a high throughput EST sequencing system using
            poly(A) tail-removed cDNA libraries and determination of 36,000
            bovine ESTs
  JOURNAL   Nucleic Acids Res. 29 (22), E108 (2001)
  MEDLINE   21570554
COMMENT     Contact: Yoshikazu Sugimoto
            Animal Genetics Division
            Shirakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
            Fax: 81-248-25-5725
            Email: kazusugi@cocoa.ocn.ne.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
FEATURES             Location/Qualifiers
     source          1. .506
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone="E1KI039B05"
                     /clone_lib="Bos taurus kidney fetus"
                     /tissue_type="kidney"
                     /dev_stage="fetus"
                     /lab_host="DH10B"
                     /note="Vector: pZL1; Site_1: Sal1; Site_2: Not1; Poly A
                     was deleted from a Not1 site"
BASE COUNT      124 a    126 c    129 g    126 t      1 others
ORIGIN
```

```
  Query Match             74.1%;  Score 438.8;  DB 10;  Length 506;
  Best Local Similarity   99.3%;  Pred. No. 3.1e-117;
  Matches  440;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy    150 TCAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACC 209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    506 TCAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACC 447

Qy    210 CTACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGT 269
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    446 CTACACCGCTGTCAAGGACTTNGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGT 387

Qy    270 TAGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGA 329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    386 TAGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGA 327

Qy    330 GAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTG 389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    326 GAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTG 267

Qy    390 GAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGA 449
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    266 GAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGA 207

Qy    450 CACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAA 509
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    206 CACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAA 147

Qy    510 TATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAG 569
          |||||||||||||| || ||||||||||||||||||||||||||||||||||||||||||
Db    146 TATTGTCAACTCTGTTGTTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAG 87

Qy    570 GTTGTCTGAATCGAAGGCCCTGG 592
          |||||||||||||||||||||||
Db     86 GTTGTCTGAATCGAAGGCCCTGG 64
```

RESULT 6
BE048177

```
LOCUS       BE048177                666 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  tz46h09.y1 NCI_CGAP_Brn52 Homo sapiens cDNA clone IMAGE:2291681 5'
            similar to SW:DHP2_HUMAN Q13609 DNASE I HOMOLOGOUS PROTEIN DHP2
            PRECURSOR ;, mRNA sequence.
ACCESSION   BE048177
VERSION     BE048177.1  GI:8365155
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 666)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Insert Length: 1048   Std Error: 0.00
            Seq primer: -40RP from Gibco
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1. .666
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2291681"
                     /clone_lib="NCI_CGAP_Brn52"
                     /tissue_type="tumor, 5 pooled (see description)"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; This library represents the normalized
                     version of NCI_CGAP_Brn35. Cloned unidirectionally.
                     Primer: Oligo dT.  Average insert size 1.19 kb. Tumor
                     types include: meningioma, oligodendroglioma, astrocytoma
                     (grade II), medulloblastoma, astrocytoma (grade IV).
```

```
                    Constructed by Life Technologies."
BASE COUNT      198 a    149 c    170 g    147 t      2 others
ORIGIN

  Query Match             73.0%;  Score 432;  DB 10;  Length 666;
  Best Local Similarity   87.5%;  Pred. No. 3.4e-115;
  Matches  471;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||||
Db     121 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 180

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db     181 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 240

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db     241 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 300

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     301 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 360

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          |||||||||||||||| |||||||||||||  ||||| |||||| ||| ||| ||||||
Db     361 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 420

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db     421 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 480

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||| ||| |||||||| ||||||| ||||||| |
Db     481 GGCTGCAGCTACGTCCCCAAGAAGGCCTGNAAGAACATCCGCTTGAGGACTGACCCCANG 540

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || ||||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db     541 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 600

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAA 538
          |||||||||| |||||||||||||||  | |  ||  ||||     ||  | ||||
Db     601 TATGACAGGATTGTGCTTAGAGGACAAAGAAATCGCAGTTCTGTGTTCCCCAAGTCAA 658


RESULT 7
BI522352/c
LOCUS       BI522352                695 bp    mRNA    linear   EST 29-AUG-2001
DEFINITION  603081420T1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5220532 3',
            mRNA sequence.
ACCESSION   BI522352
VERSION     BI522352.1  GI:15347144
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 695)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11554  row: h  column: 05
            High quality sequence start: 18
            High quality sequence stop: 695.

FEATURES             Location/Qualifiers
     source          1. .695
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5220532"
                     /clone_lib="NIH_MGC_120"
                     /lab_host="DH10B"
                     /note="Organ: pooled pancreas and spleen; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of spleen and pancreas from 28 yo
                     male.  Library is oligo-dT primed and directionally cloned
                     (EcoRV  site is destroyed upon cloning).  Average insert
                     size 1.5 kb, insert size range 1-2.5 kb.  Library is
                     normalized and enriched for  full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT      159 a    171 c    163 g    202 t
ORIGIN

  Query Match             71.1%;  Score 421;  DB 13;  Length 695;
  Best Local Similarity   87.2%;  Pred. No. 5.7e-112;
  Matches  496;  Conservative    0;  Mismatches   70;  Indels    3;  Gaps    3;

Qy      23 GAAAAGGCATAACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAG 82
           | | ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     695 GGAGAGGCATAACATACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAG 636

Qy      83 AACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACC 142
           ||| |||||||||||||| || |||||||| |||||||| ||    || || | |||||
Db     635 CACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACC 576

Qy     143 ACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCC 202
           | |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||
Db     575 ATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCC 516

Qy     203 AGTCACCCTACA-CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAG 261
           | || ||| ||| | |||||||| |||||||||||| |||||||||||||||||| |||
Db     515 AATCTCCCCACATCTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAG 456

Qy     262 ACATCCGTTAGAGAGATTGATGAGCTGGCT-GATGTCTACACAGATGTGAAACGTCGCTG 320
           |||||||||| ||||| |||||| ||| | || |||||||| || |||||||  |||||
Db     455 ACATCCGTTAAGGAGATCGATGAGTTGGTTAGAGGTCTACACGGACGTGAAACACCGCTG 396

Qy     321 GAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAA 380
           ||| || |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     395 GAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAA 336

Qy     381 GAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGA 440
           |||||||||||| |||||||| ||||||| ||||||| ||| ||||||||||||||||||
Db     335 GAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGA 276

Qy     441 CCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAG 500
           ||||||||||||||||| |||||||||| ||||| || |||||||||| ||||||||||
Db     275 CCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAG 216

Qy     501 AGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCA-AACCTCGTCTTTGATTTCCAGA 559
           ||||||| | || |||| |||| || ||| | ||| |||   || |||||  |||||||
Db     215 AGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAGAACAGTGTTTTTGACTTCCAGA 156

Qy     560 AAGCTTACAGGTTGTCTGAATCGAAGGCC 588
           ||||||||| | || |||||  | |||||
Db     155 AAGCTTACAAGCTGACTGAAGAGGAGGCC 127


RESULT 8
AW543356/c
LOCUS       AW543356                642 bp    mRNA    linear   EST 31-AUG-2000
DEFINITION  C0168H07-3 NIA Mouse E7.5 Extraembryonic Portion cDNA Library Mus
            musculus cDNA clone C0168H07 3', mRNA sequence.
ACCESSION   AW543356
VERSION     AW543356.1  GI:7185773
KEYWORDS    EST.
SOURCE      house mouse.
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
            ,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
            III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
  MEDLINE   20381348
COMMENT     Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: C0168  row: H  column: 07
            Seq primer: -21M13 Forward
            High quality sequence stop: 642
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .642
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:C0168H07-3"
                     /db_xref="taxon:10090"
                     /clone="C0168H07"
                     /clone_lib="NIA Mouse E7.5 Extraembryonic Portion cDNA
                     Library"
                     /sex="unknown"
                     /dev_stage="7.5dpc Embryo"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1 (Gibco/BRL Life Technology);
                     Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
                     5 EPC. The double-stranded cDNA was  synthesized by
                     Gibco's kit with an Oligo(dT) primer [NotI primer-adapter
                     from GibcoBRL]
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT-3']
                     from 0.8ug of mRNA . The double-stranded cDNAs were
                     treated with T4  DNA polymerase and purified by
                     ethanol-precipitation. The cDNAs were  ligated to
                     Lone-linker LL-Sal3 (include Sal1 sequence). The cDNAs
                     were  purified by phenol/chloroform and separated from
                     free linkers by  Centricon 100. Then, cDNAs were amplified
                     by long-range high fidelity  PCR using Takara's Ex Taq
                     polymerase. Then, the cDNAs were purified by
                     phenol/chloroform and by Centricon 100. The cDNAs were
                     digested with SalI  and NotI enzymes. Then, the cDNAs were
                     size selected by Gibco's Size  Fractionation Column. The
                     cDNAs were cloned into SalI/NotI site of  pSPORT1 plasmid
                     vector. The DH10B E. coli host was transformed with the
                     ligation mixture by chemical method. The library was
                     constructed by  Xiaohong Wang."
BASE COUNT      166 a    161 c    148 g    167 t
ORIGIN

  Query Match             69.5%;  Score 411.6;  DB 10;  Length 642;
  Best Local Similarity   83.3%;  Pred. No. 3.1e-109;
  Matches  468;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy      31 ATAACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTAT 90
           | |||||||||||||||||||| ||||| |||||||||||||| || ||||| ||||||
Db     642 ACAACATACAACTATGTGATTAGTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTAT 583

Qy      91 GCCTTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTAT 150
           |||||  |||| || || ||||| |||||||| ||   ||  |||| |||||| ||||||
Db     582 GCCTTCGTCTACAAGGAGAAGCTGGTGTCTGTGAAGACAAAATACCACTACCATGACTAT 523

Qy     151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210
           |||| ||||||| |||| |||||||||||||| |||||||||| |||||||| || |||
Db     522 CAGGATGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCC 463

Qy     211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270
           | ||||||||||||||||||||||||||||| ||||||| || || ||||| ||||||
Db     462 TTTACTGCTGTCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTT 403

Qy     271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330
           | |||||| |||||||||  |||||||||| ||||||| | | |  |||||  |||||
Db     402 AAAGAGATAGATGAGCTGGTCGATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAG 343

Qy     331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390
           |||||||| |||||||||| ||||| || ||||| ||||| ||||||||||||||||||
Db     342 AATTTCATCTTCATGGGTGATTTCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGG 283

Qy     391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450
           || |||| ||  |||||||||||||||||||| |||||||||| ||||||||||||||||
Db     282 CAGAACATTCGTTTGAGGACGGACCCCAAGTTTGTTTGGCTGATTGGGGACCAAGAGGAC 223

Qy     451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510
           || |||||||||||||| || | ||| |||||||||||||| |||||| | |||||| |
Db     222 ACTACGGTCAAGAAGAGTACCAGCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAG 163

Qy     511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570
           || |||||||| |  |  || |  || |    ||||||||| || ||||||||||
Db     162 ATAGTCAACTCCGTGGTCCCCCGTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGAC 103

Qy     571 TTGTCTGAATCGAAGGCCCTGG 592
           ||||||||   | |||||||||
Db     102 TTGTCTGAGGAGGAGGCCCTGG 81


RESULT 9
BG867772
LOCUS       BG867772                825 bp    mRNA    linear   EST 29-MAY-2001
DEFINITION  602786883F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912893 5',
            mRNA sequence.
ACCESSION   BG867772
VERSION     BG867772.1  GI:14218312
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 825)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10817  row: e  column: 22
            High quality sequence stop: 799.
FEATURES             Location/Qualifiers
     source          1. .825
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4912893"
                     /clone_lib="NCI_CGAP_SG2"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
                     NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
                     dT. Average insert size 1.3 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      215 a    203 c    204 g    203 t
ORIGIN

  Query Match             68.6%;  Score 406;  DB 12;  Length 825;
  Best Local Similarity   82.7%;  Pred. No. 1.5e-107;
  Matches  488;  Conservative    0;  Mismatches  100;  Indels    2;  Gaps    2;
```

```
Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| || ||||||||| ||| | ||| || |||||||||||||||||| |||||
Db     221 GAGAAGCTGAATGGAAATTCACGAAGAAGCACAATATACAACTATGTGATTAGTTCTCGA 280

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||| || ||||| |||||||||| |||| || || ||||| ||||||
Db     281 CTTGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGCTGGTGTCT 340

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||   ||  |||| |||||| |||||||||| ||||||| |||| ||||||||||||
Db     341 GTGAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 400

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||| |||||||| || |||| || ||||||||||||||||||||||||||
Db     401 GAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCGTGATTGTC 460

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||| ||||||| || || ||||| ||||||| |||||| |||||||||  |||||||||
Db     461 CCCTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCGATGTCTAC 520

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || ||||||| | | |  |||||  |||||||||||| |||||||||| ||||| ||
Db     521 ACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATTTCAACGCC 580

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGA-GGACGGACCCCAA 419
          ||||| ||||| |||||||||||||||||| || |||| ||  ||| ||||||||||||
Db     581 GGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGACGGACGGACCCCAA 640

Qy     420 GTTCG-TTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCG 478
          ||| | |||||||||| |||||||||||||||| |||||||||||||| || | ||| |
Db     641 GTTTGTTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCAGCTGTG 700

Qy     479 CCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAA 538
          |||||||||||| |||||| | |||||| | || |||||||| |  | ||| |  || |
Db     701 CCTATGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTTCCCCGTTCCA 760

Qy     539 ACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCC 588
              ||||||||| || ||||||||||    ||||||||  |  || ||
Db     761 GTGGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGACGCAGAGCCC 810


RESULT 10
BG530136
LOCUS       BG530136                825 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602558626F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4696638 5',
            mRNA sequence.
ACCESSION   BG530136
VERSION     BG530136.1  GI:13521673
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 825)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1524  row: c  column: 07
            High quality sequence stop: 743.
FEATURES             Location/Qualifiers
     source          1. .825
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4696638"
                     /clone_lib="NIH_MGC_61"
                     /tissue_type="embryonal carcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
BASE COUNT      225 a    204 c    217 g    179 t
ORIGIN

  Query Match             67.0%;  Score 396.4;  DB 12;  Length 825;
  Best Local Similarity   82.5%;  Pred. No. 9.5e-105;
  Matches  490;  Conservative    0;  Mismatches  101;  Indels    3;  Gaps    3;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||||
Db     232 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 291

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||||||||||||||||||| || |||||||| ||||||
Db     292 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 351

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGG-AGACGCAGATGTGTTTTCCAG 179
          || ||    || || | |||||| |||||||||| ||| ||||||||||||||||||||
Db     352 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGCAGACGCAGATGTGTTTTCCAG 411

Qy     180 GGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGT 239
          ||| ||||||||||||||||||| || ||| |||| |||||||| |||||||||||| |
Db     412 GGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTAT 471

Qy     240 CCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTA 299
          ||||||||||||| ||| |||||||||||||  ||||| |||||| ||| ||| |||||
Db     472 CCCCCTGCACACCAGCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTA 531

Qy     300 CACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGC 359
          ||| || |||||||  |||||||| || | ||||||||||||||||||||||||||||||
Db     532 CACGGACGTGAAACACCGCTGGAAGGCGGCGAATTTCATTTTCATGGGTGACTTCAATGC 591

Qy     360 TGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAA-GGACATCCGCCTGAGGACGGACCCCA 418
          |||||||||      |||||||||| |||||| | |||||||| ||||||| |||||||
Db     592 CGGCTGCAGCTTACGTCCCAAGAAGGTCTGGAACGAACATCCGCTTGAGGACTGACCCCA 651

Qy     419 AGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCG 478
          ||| | ||||||||||||||||||||||||||||||||||||| |||||||||| |||| |
Db     652 GGTTTGCTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTG 711

Qy     479 CCTATGACAGGA-TCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCA 537
          | |||||||||| | |||||||  | || || |      | |      | |||
Db     712 CATATGACAGGATTGGTGCTTAGCAGGCACAAGAAATCGTCCAGTCTGTCGTCCCAAGTC 771

Qy     538 AACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTG 591
          || |  |   |||| |||||||||||| |||| || ||| |    ||||| ||
Db     772 AAACAGGGTCTTGACTTCCAGAAAGCTTCCAGGCTGACTGTAGAAGAGGCCTTG 825


RESULT 11
AW543802/c
LOCUS       AW543802                620 bp    mRNA    linear   EST 31-AUG-2000
DEFINITION  C0174C07-3 NIA Mouse E7.5 Extraembryonic Portion cDNA Library Mus
            musculus cDNA clone C0174C07 3', mRNA sequence.
ACCESSION   AW543802
VERSION     AW543802.1  GI:7186219
KEYWORDS    EST.
SOURCE      house mouse.
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 620)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
            ,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
            III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
  MEDLINE   20381348
COMMENT     Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: C0174  row: C  column: 07
            Seq primer: -21M13 Forward
            High quality sequence stop: 620
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .620
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:C0174C07-3"
                     /db_xref="taxon:10090"
                     /clone="C0174C07"
                     /clone_lib="NIA Mouse E7.5 Extraembryonic Portion cDNA
                     Library"
                     /sex="unknown"
                     /dev_stage="7.5dpc Embryo"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1 (Gibco/BRL Life Technology);
                     Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
                     5 EPC. The double-stranded cDNA was  synthesized by
                     Gibco's kit with an Oligo(dT) primer [NotI primer-adapter
                     from GibcoBRL]
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT-3']
                     from 0.8ug of mRNA . The double-stranded cDNAs were
                     treated with T4  DNA polymerase and purified by
                     ethanol-precipitation. The cDNAs were  ligated to
                     Lone-linker LL-Sal3 (include Sal1 sequence). The cDNAs
                     were  purified by phenol/chloroform and separated from
                     free linkers by  Centricon 100. Then, cDNAs were amplified
                     by long-range high fidelity  PCR using Takara's Ex Taq
                     polymerase. Then, the cDNAs were purified by
                     phenol/chloroform and by Centricon 100. The cDNAs were
                     digested with SalI  and NotI enzymes. Then, the cDNAs were
                     size selected by Gibco's Size  Fractionation Column. The
                     cDNAs were cloned into SalI/NotI site of  pSPORT1 plasmid
                     vector. The DH10B E. coli host was transformed with the
                     ligation mixture by chemical method. The library was
                     constructed by  Xiaohong Wang."
BASE COUNT      160 a    159 c    143 g    158 t
ORIGIN

  Query Match           66.1%;  Score 391.2;  DB 10;  Length 620;
  Best Local Similarity  82.8%;  Pred. No. 2.8e-103;
  Matches  447;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy      53 GCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGC 112
           | ||||| ||||||||||||||| || ||||| |||||||||| |||| || || ||||
Db     620 GTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGC 561

Qy     113 TAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGT 172
           | |||||||| ||   ||  |||| |||||| |||||||||| ||||||| |||| ||||
Db     560 TGGTGTCTGTGAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGT 501

Qy     173 TTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCG 232
           |||||||||| |||||||||| |||||||| || ||||  || ||||||||||||||||
Db     500 TTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCG 441

Qy     233 TGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTG 292
           |||||||||| ||||||| || || ||||| ||||||| ||||||  |||||||||| |
Db     440 TGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCG 381

Qy     293 ATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACT 352
           |||||||||| ||||||| | | |  |||||  |||||||||||| |||||||||| |
Db     380 ATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATT 321

Qy     353 TCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGG 412
           |||| || ||||| ||||| |||||||||||||||||| || |||| ||  |||||||||
Db     320 TCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGAGGACGG 261

Qy     413 ACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAA 472
           |||||||||| |||||||||| |||||||||||||||||| |||||||||||||| || |
Db     260 ACCCCAAGTTTGTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCA 201

Qy     473 ACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTC 532
           ||| |||||||||||||| |||||| | |||||| | || |||||||| |  |  || |
Db     200 GCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTACCCC 141

Qy     533 AATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
             || |    ||||||||| || |||||||||||    ||||||||   | |||||||||
Db     140 GTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 81


RESULT 12
BI460950
LOCUS       BI460950                628 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603207187F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5272917 5',
            mRNA sequence.
ACCESSION   BI460950
VERSION     BI460950.1  GI:15251606
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 628)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11688  row: n  column: 22
            High quality sequence stop: 628.
FEATURES             Location/Qualifiers
     source          1. .628
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5272917"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      181 a    144 c    166 g    137 t
ORIGIN

  Query Match           65.8%;  Score 389.4;  DB 13;  Length 628;
  Best Local Similarity  90.1%;  Pred. No. 9.4e-103;
```

```
  Matches  417;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||||||||||||||||||||||||||||
Db     166 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACATACAACTATGTGATTAGCTCTCGG 225

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||||| |||||||||||||| || ||||||||| ||||||
Db     226 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 285

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db     286 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 345

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     346 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 405

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          |||||||||||||||| |||||||||||| ||||| |||||| ||| ||| ||||||
Db     406 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 465

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db     466 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 525

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db     526 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 585

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGA 463
          || ||||||||||||||||||||||||||||||||| ||||
Db     586 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGA 628


RESULT 13
AW540389/c
LOCUS       AW540389                610 bp    mRNA    linear   EST 31-AUG-2000
DEFINITION  C0132C04-3 NIA Mouse E7.5 Extraembryonic Portion cDNA Library Mus
            musculus cDNA clone C0132C04 3', mRNA sequence.
ACCESSION   AW540389
VERSION     AW540389.1  GI:7182806
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 610)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
            ,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
            III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
  MEDLINE   20381348
COMMENT     Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: C0132  row: C  column: 04
            Seq primer: -21M13 Forward
            High quality sequence stop: 610
            POLYA=Yes.
FEATURES            Location/Qualifiers
     source         1. .610
                    /organism="Mus musculus"
                    /strain="C57BL/6J"
                    /db_xref="niaEST:C0132C04-3"
                    /db_xref="taxon:10090"
                    /clone="C0132C04"
                    /clone_lib="NIA Mouse E7.5 Extraembryonic Portion cDNA
                    Library"
                    /sex="unknown"
                    /dev_stage="7.5dpc Embryo"
                    /lab_host="DH10B"
                    /note="Vector: pSPORT1 (Gibco/BRL Life Technology);
                    Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
                    5 EPC. The double-stranded cDNA was  synthesized by
                    Gibco's kit with an Oligo(dT) primer [NotI primer-adapter
                    from GibcoBRL]
                    [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT-3']
                    from 0.8ug of mRNA . The double-stranded cDNAs were
                    treated with T4  DNA polymerase and purified by
                    ethanol-precipitation. The cDNAs were  ligated to
                    Lone-linker LL-Sal3 (include Sal1 sequence). The cDNAs
                    were  purified by phenol/chloroform and separated from
                    free linkers by  Centricon 100. Then, cDNAs were amplified
                    by long-range high fidelity  PCR using Takara's Ex Taq
                    polymerase. Then, the cDNAs were purified by
                    phenol/chloroform and by Centricon 100. The cDNAs were
                    digested with SalI  and NotI enzymes. Then, the cDNAs were
                    size selected by Gibco's Size  Fractionation Column. The
                    cDNAs were cloned into SalI/NotI site of  pSPORT1 plasmid
                    vector. The DH10B E. coli host was transformed with the
                    ligation mixture by chemical method. The library was
                    constructed by  Xiaohong Wang."
BASE COUNT      157 a    157 c    140 g    156 t
ORIGIN

  Query Match           65.2%;  Score 386;  DB 10;  Length 610;
  Best Local Similarity   83.0%;  Pred. No. 9.1e-102;
  Matches  440;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy      63 TGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCTGT 122
           |||||||||||| || ||||| ||||||||||  |||| || || ||||| ||||||||
Db     610 TGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGCTGGTGTCTGT 551

Qy     123 AAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGGGA 182
           ||   ||  |||| |||||| |||||||||| ||||||| |||| ||||||||||||||
Db     550 GAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGGGA 491

Qy     183 ACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTCCC 242
           |||||||||| |||||||| || ||||  || ||||||||||||||||||||||||||||
Db     490 GCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCGTGATTGTCCC 431

Qy     243 CCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTACAC 302
           | ||||||| || || ||||| ||||||| |||||| ||||||||||  ||||||||||
Db     430 CTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCGATGTCTACAC 371

Qy     303 AGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCTGG 362
           ||||||| | | |  |||||  |||||||||||||| |||||||||| ||||| || ||
Db     370 GGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATTTCAACGCCGG 311

Qy     363 CTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTT 422
           ||| |||||||||||||||||||||| || |||| ||  ||||||||||||||||||||
Db     310 CTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGAGGACGGACCCCAAGTT 251

Qy     423 CGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCCTA 482
           |||||||||| |||||||||||||||||| |||||||||||||| || | ||| |||||
Db     250 TGTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCAGCTGTGCCTA 191

Qy     483 TGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAACCT 542
           ||||||||| |||||| | |||||| | || |||||||| |  | ||| |  || |
Db     190 TGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTTCCCCGTTCCAGTGG 131

Qy     543 CGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
           ||||||||| || |||||||||||     |||||||   | |||||||||
Db     130 CGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 81


RESULT 14
AW915564
LOCUS       AW915564                668 bp    mRNA    linear   EST 25-MAY-2000
DEFINITION  EST346868 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
            RGICW80 5' end, mRNA sequence.
```

```
ACCESSION   AW915564
VERSION     AW915564.1  GI:8081263
KEYWORDS    EST.
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
            Kerlavage,A.R. and Adams,M.D.
  TITLE     Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
            Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Lee, NH
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Tel: (301)-838-3529
            Fax: (301)-838-0208
            Email: nhlee@tigr.org
            This clone is available through the ATCC, contact the ATCC
            tel#703-365-2700 for further information
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="Rattus sp."
                     /db_xref="taxon:10118"
                     /clone="RGICW80"
                     /clone_lib="Normalized rat embryo, Bento Soares"
                     /dev_stage="embryo 8, 12, 18 dpc"
                     /note="Vector: pT7T3Pac; Site_1: EcoRI; Site_2: NotI"
BASE COUNT      172 a    160 c    168 g    168 t
ORIGIN

  Query Match           64.8%;  Score 383.8;  DB 10;  Length 668;
  Best Local Similarity  85.6%;  Pred. No. 4.2e-101;
  Matches  439;  Conservative    0;  Mismatches   72;  Indels    2;  Gaps    1;

Qy      80 AAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCT 139
           |||||||||||||||| ||||| || || ||||| |||||||| ||    ||   ||||||
Db       2 AAGAACAGTATGCCTTCCTCTACAAGGAGAAGCTGGTGTCTGTGAAGGCAAA--ACCTCT 59

Qy     140 ACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGT 199
           |||| |||||||||| ||||||| |||| |||||||||||||| |||||||||| ||||
Db      60 ACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGT 119

Qy     200 TCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTG 259
           ||||| | |||| ||| |||| |||||||||||||||||||||| ||||||| || ||||
Db     120 TCCAGGCGCCCTTCACTGCTGCCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCTG 179

Qy     260 AGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCT 319
           | || ||||||| |||||| |||||||||||||| |||||||| ||||||| | | || |
Db     180 AAACCTCCGTTAAAGAGATAGATGAGCTGGCTGACGTCTACACGGATGTGAGAAGACGAT 239

Qy     320 GGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCA 379
           |||| |||||||||||||| |||||||||| ||||||||||||||||||||||||||||
Db     240 GGAAGGCAGAGAATTTCATCTTCATGGGTGATTTCAATGCTGGCTGCAGCTACGTCCCCA 299

Qy     380 AGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGG 439
           |||||||||||||| ||||||  ||||||| |||||||| || |||||||||| ||||
Db     300 AGAAGGCCTGGAAGAACATCCGTTTGAGGACAGACCCCAACTTTGTTTGGCTGATTGGGG 359

Qy     440 ACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTA 499
           |||||||||||||||||||||||||||||| | ||| |||||||||||||| ||||||
Db     360 ACCAAGAGGACACCACGGTCAAGAAGAGCACCAGCTGTGCCTATGACAGGATTGTGCTTC 419

Qy     500 GAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGA 559
           | |||||| | || |||||||||  | ||| |  || |    |||||||| || ||||
Db     420 GCGGACAAGAGATAGTCAACTCTGTGGTTCCCCGCTCCAGTGGCGTCTTTGACTTTCAGA 479

Qy     560 AAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
           |||||||   |||||||||  | |||||||||
Db     480 AAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGG 512
```

```
RESULT 15
AI927844/c
LOCUS       AI927844                784 bp    mRNA    linear   EST 08-MAR-2000
DEFINITION  wo91g09.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2462752 3'
            similar to SW:DHP2_HUMAN Q13609 DNASE I HOMOLOGOUS PROTEIN DHP2
            PRECURSOR ;, mRNA sequence.
ACCESSION   AI927844
VERSION     AI927844.1  GI:5663808
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 739   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 426.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2462752"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
BASE COUNT      173 a    188 c    179 g    242 t      2 others
ORIGIN

  Query Match           64.8%;  Score 383.6;  DB 9;  Length 784;
  Best Local Similarity  86.7%;  Pred. No. 5.1e-101;
  Matches  422;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy     106 GAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCA 165
           |||||||| |||||||| ||    || || | |||||| |||||||||| ||||||||||
Db     667 GAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCA 608

Qy     166 GATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAG 225
           |||||||||||||||| |||| |||||||||||||| || ||| |||| ||||||||
Db     607 GATGTGTTTTCCAGGGAGCCCTGTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAA 548

Qy     226 GACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAG 285
           |||||||||||| |||||||||||||||| | ||||||||||||  |||||  ||||||
Db     547 GACTTCGTGATTATCCCCCTGCACACCACCNCAGAGACATCCGTTAAGGAGATCGATGAG 488

Qy     286 CTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATG 345
           ||| ||| |||||||| || |||||||  |||||||| || ||||||||||||||||||
Db     487 TTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATG 428
```

```
Qy    346 GGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTG 405
          ||||||||||||| |||||||||||||||||||||||||||||||| |||||||| ||
Db    427 GGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTG 368

Qy    406 AGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAG 465
          ||||| ||||||| ||| ||||||||||||||||||||||||||||||||||| ||||||
Db    367 AGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAG 308

Qy    466 AGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGT 525
          ||||| ||||| || |||||||||| |||||||||||||| | || ||||  |||| |
Db    307 AGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTT 248

Qy    526 GGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAG 585
          | |||  | ||||||   || ||||| |||||||||||||||| | || |||||  | ||
Db    247 GTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAG 188

Qy    586 GCCCTGG 592
          |||||||
Db    187 GCCCTGG 181
```

Search completed: December 9, 2002, 12:18:03
Job time : 1339.4 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 08:59:53 ; Search time 36.0918 Seconds
(without alignments)
5030.295 Million cell updates/sec

Title: US-09-905-114-1
Perfect score: 592
Sequence: 1 gagaagctaaacggaaattc.....gtctgaatcgaaggccctgg 592

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	452.8	76.5	1208	1	US-08-640-765A-2	Sequence 2, Appli
	2	452.8	76.5	1208	3	US-09-073-613-2	Sequence 2, Appli
	3	149.4	25.2	1039	1	US-07-895-300A-2	Sequence 2, Appli
	4	149.4	25.2	1039	4	US-09-638-112-2	Sequence 2, Appli
	5	149.4	25.2	1039	5	PCT-US93-05136-2	Sequence 2, Appli
	6	149.4	25.2	1039	5	PCT-US93-10519-1	Sequence 1, Appli
	7	147.8	25.0	1039	1	US-08-458-367-2	Sequence 2, Appli
	8	117.2	19.8	1615	2	US-08-491-988-15	Sequence 15, Appl
	9	117.2	19.8	1624	2	US-08-491-988-14	Sequence 14, Appl
	10	117.2	19.8	1648	2	US-08-491-988-13	Sequence 13, Appl
	11	115.2	19.5	1055	2	US-08-468-012A-1	Sequence 1, Appli
	12	115.2	19.5	1055	4	US-09-054-989-1	Sequence 1, Appli
	13	35.2	5.9	1427	3	US-09-173-581-10	Sequence 10, Appl
	14	35.2	5.9	1427	4	US-09-420-915-10	Sequence 10, Appl
	15	32.8	5.5	1293	2	US-08-924-440-1	Sequence 1, Appli
	16	32.2	5.4	289	4	US-09-007-005-17	Sequence 17, Appl
	17	32.2	5.4	289	4	US-09-244-796-17	Sequence 17, Appl
c	18	32.2	5.4	7218	1	US-08-232-463-14	Sequence 14, Appl
	19	31.8	5.4	1511	2	US-08-809-763-3	Sequence 3, Appli
	20	31.8	5.4	1511	3	US-08-956-253-3	Sequence 3, Appli
c	21	31.6	5.3	1001	4	US-09-641-638-90	Sequence 90, Appl
c	22	31.2	5.3	248	3	US-08-513-974B-300	Sequence 300, App
c	23	31.2	5.3	248	3	US-08-513-974B-301	Sequence 301, App
	24	30.6	5.2	1920	1	US-08-487-753-1	Sequence 1, Appli
	25	30.6	5.2	1920	2	US-08-480-065-1	Sequence 1, Appli
	26	30.6	5.2	1920	3	US-08-487-744-1	Sequence 1, Appli
	27	30.6	5.2	1920	5	PCT-US93-09167-1	Sequence 1, Appli
c	28	30.6	5.2	2900	2	US-09-027-337-9	Sequence 9, Appli
c	29	30.6	5.2	2900	4	US-09-644-600-9	Sequence 9, Appli
c	30	30.6	5.2	3147	2	US-09-027-337-1	Sequence 1, Appli
c	31	30.6	5.2	3147	4	US-09-644-600-1	Sequence 1, Appli
	32	30.6	5.2	3147	4	US-09-644-600-18	Sequence 18, Appl
c	33	30	5.1	1734	4	US-09-484-970B-63	Sequence 63, Appl
	34	30	5.1	2071	1	US-07-923-724-1	Sequence 1, Appli
	35	30	5.1	2071	2	US-08-609-426A-1	Sequence 1, Appli
	36	30	5.1	2071	2	US-08-374-652C-3	Sequence 3, Appli
c	37	30	5.1	4049	1	US-08-162-809-17	Sequence 17, Appl
c	38	30	5.1	4097	1	US-08-162-809-11	Sequence 11, Appl
c	39	29.6	5.0	352	4	US-09-228-986-21	Sequence 21, Appl
c	40	29.6	5.0	560	3	US-09-042-426-5	Sequence 5, Appli
c	41	29.6	5.0	560	4	US-09-291-238-5	Sequence 5, Appli
c	42	29.6	5.0	560	4	US-09-330-760-5	Sequence 5, Appli
c	43	29.6	5.0	560	4	US-09-328-473-5	Sequence 5, Appli
c	44	29.6	5.0	560	4	US-09-330-737-5	Sequence 5, Appli
c	45	29.6	5.0	560	4	US-09-329-169-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-08-640-765A-2
; Sequence 2, Application US/08640765A
; Patent No. 5821103
;  GENERAL INFORMATION:
;    APPLICANT:  Tanuma, Sei-ichi
;    TITLE OF INVENTION:  NOVEL DEOXYRIBONUCLEASE
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/640,765A
;      FILING DATE:  06-MAY-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6239518
;      FILING DATE:  06-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Green, Robert F.
;      REGISTRATION NUMBER:  27555
;      REFERENCE/DOCKET NUMBER:  73362
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5600
;      TELEFAX:  (312) 616-5700
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1208 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  12..941
US-08-640-765A-2

  Query Match           76.5%;  Score 452.8;  DB 1;  Length 1208;
  Best Local Similarity  85.3%;  Pred. No. 5.4e-132;
  Matches  505;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;
```

```
Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
Db     243 GAGAAGCTGAATGGAAACTCACGAAGAAGCACGACATACAACTACGTGATTAGCTCTCGG 302

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
Db     303 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTCCTCTACAAGGAGAAGCTGGTGTCT 362

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
Db     363 GTGAAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 422

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
Db     423 GAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCACTGCTGCCAAGGACTTCGTGATTGTC 482

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
Db     483 CCCTTGCACACAACTCCTGAAACCTCCGTTAAAGAGATAGATGAGCTGGCTGACGTCTAC 542

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
Db     543 ACGGATGTTAGAAGACGATGGAAGGCAGAGATTTTCATCTTCATGGGTGATTTCAATGCT 602

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
Db     603 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGTTTGAGGACAGACCCCAAC 662

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
Db     663 TTTGTTTGGCTGATTGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACCAGCTGTGCC 722

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
Db     723 TATGACAGGATTGTCGTTCGCGGACAAGAGATAGTCAACTCTGTGGTTCCCCGCTCCAGT 782

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
Db     783 GGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGG 834
```

RESULT 2
US-09-073-613-2
; Sequence 2, Application US/09073613
; Patent No. 6143875
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; TITLE OF INVENTION: ANTIBODY TO NOVEL DEOXYRIBONUCLEASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,613
; FILING DATE: 06-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/640,765
; FILING DATE:
; APPLICATION NUMBER: JP 6239518
; FILING DATE: 06-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: LARCHER, Carol
; REGISTRATION NUMBER: 35243
; REFERENCE/DOCKET NUMBER: 85519
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1208 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12..941
US-09-073-613-2

Query Match 76.5%; Score 452.8; DB 3; Length 1208;
Best Local Similarity 85.3%; Pred. No. 5.4e-132;
Matches 505; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```
Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
Db     243 GAGAAGCTGAATGGAAACTCACGAAGAAGCACGACATACAACTACGTGATTAGCTCTCGG 302

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
Db     303 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTCCTCTACAAGGAGAAGCTGGTGTCT 362

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
Db     363 GTGAAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 422

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
Db     423 GAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCACTGCTGCCAAGGACTTCGTGATTGTC 482

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
Db     483 CCCTTGCACACAACTCCTGAAACCTCCGTTAAAGAGATAGATGAGCTGGCTGACGTCTAC 542

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
Db     543 ACGGATGTTAGAAGACGATGGAAGGCAGAGATTTTCATCTTCATGGGTGATTTCAATGCT 602

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
Db     603 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGTTTGAGGACAGACCCCAAC 662

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
Db     663 TTTGTTTGGCTGATTGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACCAGCTGTGCC 722

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
Db     723 TATGACAGGATTGTCGTTCGCGGACAAGAGATAGTCAACTCTGTGGTTCCCCGCTCCAGT 782

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
Db     783 GGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGG 834
```

RESULT 3
US-07-895-300A-2
; Sequence 2, Application US/07895300A
; Patent No. 5279823
; GENERAL INFORMATION:
; APPLICANT: Frenz, John
; APPLICANT: Shire, Steven J.
; APPLICANT: Sliwkowski, Mary B.
; TITLE OF INVENTION: PURIFIED FORMS OF DNase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:

```
;       ADDRESSEE:  Genentech, Inc.
;       STREET:  460 Point San Bruno Blvd
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  patin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/895,300A
;       FILING DATE:  19920608
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hensley, Max D.
;       REGISTRATION NUMBER:  27,043
;       REFERENCE/DOCKET NUMBER:  747
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-1994
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1039 bases
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-07-895-300A-2

  Query Match             25.2%;  Score 149.4;  DB 1;  Length 1039;
  Best Local Similarity   56.7%;  Pred. No. 4.4e-37;
  Matches  317;  Conservative    0;  Mismatches  236;  Indels    6;  Gaps    2;

Qy      34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy      94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 150
           ||  | || |       | |  |||||||       | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210
             || |  | |      ||    ||   | | || ||     | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy     211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270
           | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |    | ||
Db     589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy     271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330
           |||||  || |  ||   ||| ||||||   |||||  ||      |||      |||
Db     649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy     331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390
           |  |||| || ||||| |||||||||||| |||||||||| ||             |||
Db     709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy     391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450
           ||||||||||  ||||   |||||  |||  |||||||||   |||   |  |||
Db     769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy     451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510
           |||||  |     |   |||  |||| |||||||||||||||||  ||  |||
Db     829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy     511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570
           |      | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db     886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy     571 TTGTCTGAATCGAAGGCCC 589
           ||  |||      |||||
Db     946 CTGAGTGACCAACTGGCCC 964


RESULT 4
US-09-638-112-2
; Sequence 2, Application US/09638112
; Patent No. 6440412
;  GENERAL INFORMATION:
;    APPLICANT:  Frenz, John
;    APPLICANT:  Shire, Steven J.
;    APPLICANT:  Sliwkowski, Mary B.
;    TITLE OF INVENTION:  PURIFIED FORMS OF DNase
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:.
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  1 DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/638,112
;      FILING DATE:  09-Aug-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/942561
;      FILING DATE:  01-OCT-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Johnston, Sean A.
;      REGISTRATION NUMBER:  35,910
;      REFERENCE/DOCKET NUMBER:  P0747C7
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650/225-3562
;      TELEFAX:  650/952-9881
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1039 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-09-638-112-2

  Query Match             25.2%;  Score 149.4;  DB 4;  Length 1039;
  Best Local Similarity   56.7%;  Pred. No. 4.4e-37;
  Matches  317;  Conservative    0;  Mismatches  236;  Indels    6;  Gaps    2;

Qy      34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy      94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 150
           ||  | || |       | |  |||||||       | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210
             || |  | |      ||    ||   | | || ||     | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy     211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270
           | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |    | ||
Db     589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy     271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330
           |||||  || |  ||   ||| ||||||   |||||  ||      |||      |||
Db     649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708
```

```
Qy    331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390
          |  |||| || |||| |||||||||| ||||||||||| ||                 |||
Db    709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy    391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450
          |||||||||| ||||   |||||  |||   |||||||||   |||   |  |||
Db    769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy    451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510
          |||||  |     |   |||  |||| ||||||||||||||||||  ||  |||
Db    829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy    511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570
          |       | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db    886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy    571 TTGTCTGAATCGAAGGCCC 589
          ||  |||      |||||
Db    946 CTGAGTGACCAACTGGCCC 964


RESULT 5
PCT-US93-05136-2
; Sequence 2, Application PC/TUS9305136
;  GENERAL INFORMATION:
;    APPLICANT:  Genentech, Inc.
;    TITLE OF INVENTION:  PURIFIED FORMS OF DNase
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/05136
;      FILING DATE:  19930528
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Johnston, Sean A.
;      REGISTRATION NUMBER:  P35,910
;      REFERENCE/DOCKET NUMBER:  747PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3562
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1039 bases
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US93-05136-2

  Query Match             25.2%;  Score 149.4;  DB 5;  Length 1039;
  Best Local Similarity   56.7%;  Pred. No. 4.4e-37;
  Matches  317;  Conservative    0;  Mismatches  236;  Indels    6;  Gaps    2;

Qy      34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy      94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 150
           ||  | || |       | |  |||||||      | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210
           || |  | |      ||     ||    | | || ||      | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy     211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270
           | ||| |  |||| ||| || |  ||||| |||||||  |  |||| | |    | ||
Db     589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy     271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330
           |||||  || |  ||   ||| ||||||   |||||  ||       |||      |||
Db     649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy     331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390
           |  |||| || ||||| |||||||||| ||||||||||| ||                 |||
Db     709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy     391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450
           |||||||||| ||||   |||||  |||   |||||||||   |||   |  |||
Db     769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy     451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510
           |||||  |     |   |||  |||| ||||||||||||||||||  ||  |||
Db     829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy     511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570
           |       | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db     886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy     571 TTGTCTGAATCGAAGGCCC 589
           ||  |||      |||||
Db     946 CTGAGTGACCAACTGGCCC 964


RESULT 6
PCT-US93-10519-1
; Sequence 1, Application PC/TUS9310519
;  GENERAL INFORMATION:
;    APPLICANT:  Genentech, Inc.
;    TITLE OF INVENTION:  Compaction Assay for Assessment of Respiratory Disease
;    TITLE OF INVENTION:  Therapy
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/10519
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Johnston, Sean A.
;      REGISTRATION NUMBER:  35,910
;      REFERENCE/DOCKET NUMBER:  792C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3562
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
```

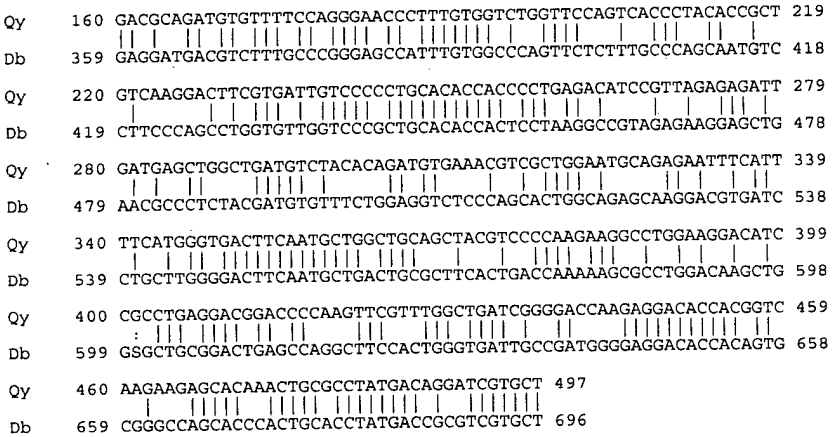
```
;      LENGTH:  1039 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US93-10519-1

  Query Match            25.2%;  Score 149.4;  DB 5;  Length 1039;
  Best Local Similarity   56.7%;  Pred. No. 4.4e-37;
  Matches  317;  Conservative    0;  Mismatches  236;  Indels    6;  Gaps    2;

Qy      34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy      94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 150
           ||  | || |      | |  |||||||      | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210
           || |  | |      ||    ||   | | || ||     | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy     211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270
           | ||| |  |||| ||| || |  ||||| |||||||| |  |||| | |    | ||
Db     589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy     271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330
           |||||  || |  ||   ||| ||||||  |||||  ||      |||      |||
Db     649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy     331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390
           |  |||| || ||||| |||||||||| ||||||||||| ||               |||
Db     709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy     391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450
           ||||||||||  ||||   |||||  |||   |||||||||   |||   |  |||
Db     769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy     451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510
           |||||  |     |    |||  |||| ||||||||||||||||||  ||  |||
Db     829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy     511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570
           |      | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db     886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy     571 TTGTCTGAATCGAAGGCCC 589
           ||  |||      |||||
Db     946 CTGAGTGACCAACTGGCCC 964


RESULT 7
US-08-458-367-2
; Sequence 2, Application US/08458367
; Patent No. 5783433
;  GENERAL INFORMATION:
;    APPLICANT:  Frenz, John
;    APPLICANT:  Shire, Steven J.
;    APPLICANT:  Sliwkowski, Mary B.
;    TITLE OF INVENTION:  PURIFIED FORMS OF DNase
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,367
;      FILING DATE:  02-Jun-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/409631
;      FILING DATE:  22-Mar-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/348284
;      FILING DATE:  30-No. 5783433-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/116186
;      FILING DATE:  02-Sep-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/895300
;      FILING DATE:  08-Jun-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Johnston, Sean A.
;      REGISTRATION NUMBER:  35,910
;      REFERENCE/DOCKET NUMBER:  P0747C4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3562
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1039 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-08-458-367-2

  Query Match            25.0%;  Score 147.8;  DB 1;  Length 1039;
  Best Local Similarity   56.5%;  Pred. No. 1.4e-36;
  Matches  316;  Conservative    0;  Mismatches  237;  Indels    6;  Gaps    2;

Qy      34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy      94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 150
           ||  | || |      | |  |||||||      | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210
           || |  | |      ||    ||   | | || ||     | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy     211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270
           | ||| |  |||| ||| || |  ||||| |||||||| |  |||| | |    | ||
Db     589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACCGAGTA 648

Qy     271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330
           |||||  || |  ||   ||| ||||||  |||||  ||      |||      |||
Db     649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy     331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390
           |  |||| ||.||||| |||||||||| |||||||||| ||               |||
Db     709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy     391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450
           ||||||||||  ||||   |||||  |||   |||||||||   |||   |  |||
Db     769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy     451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510
           |||||  |     |    |||  |||| ||||||||||||||||||  ||  |||
Db     829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy     511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570
           |      | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db     886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy     571 TTGTCTGAATCGAAGGCCC 589
```

```
Db    946 CTGAGTGACCAACTGGCCC 964


RESULT 8
US-08-491-988-15
; Sequence 15, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1615 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-15

  Query Match             19.8%;  Score 117.2;  DB 2;  Length 1615;
  Best Local Similarity   53.0%;  Pred. No. 6.5e-27;
  Matches  298;  Conservative    0;  Mismatches  258;  Indels    6;  Gaps    2;

Qy      34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93
           || ||| |||| ||  ||      | ||| ||  | ||| | || ||||||| |||
Db    1009 ACCTACCACTACGTTGTTTCTGAACCCCTCGGGCGTAACTCTTACAAAGAACGGTACCTG 1068

Qy      94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGAC---TAT 150
           || || |         ||  ||| || |||    | | |||||  ||| |||||   |
Db    1069 TTCCTGTTCCGTCCGAACAAAGTTTCAGTACTGGATACCTACCAGTACGACGACGGATGC 1128

Qy     151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210
           |  || | |      ||     || ||| |||||||   ||| ||    ||| |   | ||
Db    1129 GAATCTTGCGGTAACGACTCTTTCTCCCGGGAACCGGCTGTTGTTAAATTCTCGAGCCAC 1188

Qy     211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270
           |  |||    || || || ||||  || ||  | ||||||| | | ||         | |||
Db    1189 TCTACCAAGGTTAAAGAGTTCGCTATCGTTGCTCTGCACAGCGCGCCGTCTGACGCTGTT 1248

Qy     271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330
             || ||  |   |||   || || |||   || ||  | |     ||| |     |
Db    1249 GCTGAAATCAACTCTCTGTACGACGTTTACCTGGACGTTCAGCAGAAATGGCACCTGAAC 1308

Qy     331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390
           |  ||||  | |||||||||||||| |||| ||||    || ||  ||          |||
Db    1309 GACGTCATGCTGATGGGTGACTTCAACGCTGACTGCTCTTATGTAACCTCTTCTCAGTGG 1368

Qy     391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450
                 || || ||| | ||      | |  |||   |||||||| ||||   |  |||
Db    1369 TCATCGATTCGTCTGCGCACCTCGTCGACCTTCCAGTGGCTGATCCCGGACTCCGCTGAC 1428

Qy     451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510
           |||||  |  |      || || |||||||| || ||| | |||||  ||   |||
Db    1429 ACCACCGCTACT---AGTACCAACTGCGCTTACGACCGTATCGTTGTTGCTGGATCCCTG 1485

Qy     511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570
           |     |||| ||  ||                |   || || ||||||   || ||  |
Db    1486 CTGCAGTCTTCTGTTGTACCGGGTAGCGCGGCCCCGTTCGACTTCCAGGCTGCATATGGT 1545

Qy     571 TTGTCTGAATCGAAGGCCCTGG 592
           | ||  |    | ||| ||||
Db    1546 CTTTCGAACGAAATGGCGCTGG 1567


RESULT 9
US-08-491-988-14
; Sequence 14, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1624 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-14

  Query Match             19.8%;  Score 117.2;  DB 2;  Length 1624;
  Best Local Similarity   53.0%;  Pred. No. 6.6e-27;
  Matches  298;  Conservative    0;  Mismatches  258;  Indels    6;  Gaps    2;

Qy      34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93
           || ||| |||| ||  ||      | ||| ||  | ||| | || ||||||| |||
Db    1018 ACCTACCACTACGTTGTTTCTGAACCCCTCGGGCGTAACTCTTACAAAGAACGGTACCTG 1077

Qy      94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGAC---TAT 150
           || || |         ||  ||| || |||    | | |||||  ||| |||||   |
Db    1078 TTCCTGTTCCGTCCGAACAAAGTTTCAGTACTGGATACCTACCAGTACGACGACGGATGC 1137

Qy     151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210
```

```
Db   1138 GAATCTTGCGGTAACGACTCTTTCTCCCGGGAACCGGCTGTTGTTAAATTCTCGAGCCAC 1197

Qy    211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270

Db   1198 TCTACCAAGGTTAAAGAGTTCGCTATCGTTGCTCTGCACAGCGCGCCGTCTGACGCTGTT 1257

Qy    271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330

Db   1258 GCTGAAATCAACTCTCTGTACGACGTTTACCTGGACGTTCAGCAGAAATGGCACCTGAAC 1317

Qy    331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390

Db   1318 GACGTCATGCTGATGGGTGACTTCAACGCTGACTGCTCTTATGTAACCTCTTCTCAGTGG 1377

Qy    391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450

Db   1378 TCATCGATTCGTCTGCGCACCTCGTCGACCTTCCAGTGGCTGATCCCGGACTCCGCTGAC 1437

Qy    451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510

Db   1438 ACCACCGCTACT---AGTACCAACTGCGCTTACGACCGTATCGTTGTTGCTGGATCCCTG 1494

Qy    511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570

Db   1495 CTGCAGTCTTCTGTTGTACCGGGTAGCGCGGCCCCGTTCGACTTCCAGGCTGCATATGGT 1554

Qy    571 TTGTCTGAATCGAAGGCCCTGG 592

Db   1555 CTTTCGAACGAAATGGCGCTGG 1576


RESULT 10
US-08-491-988-13
; Sequence 13, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1648 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-13

  Query Match           19.8%;  Score 117.2;  DB 2;  Length 1648;
  Best Local Similarity   53.0%;  Pred. No. 6.6e-27;
  Matches  298;  Conservative    0;  Mismatches  258;  Indels    6;  Gaps    2;

Qy     34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93

Db   1042 ACCTACCACTACGTTGTTTCTGAACCCCTCGGGCGTAACTCTTACAAAGAACGGTACCTG 1101

Qy     94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGAC---TAT 150

Db   1102 TTCCTGTTCCGTCCGAACAAAGTTTCAGTACTGGATACCTACCAGTACGACGACGGATGC 1161

Qy    151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210

Db   1162 GAATCTTGCGGTAACGACTCTTTCTCCCGGGAACCGGCTGTTGTTAAATTCTCGAGCCAC 1221

Qy    211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270

Db   1222 TCTACCAAGGTTAAAGAGTTCGCTATCGTTGCTCTGCACAGCGCGCCGTCTGACGCTGTT 1281

Qy    271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330

Db   1282 GCTGAAATCAACTCTCTGTACGACGTTTACCTGGACGTTCAGCAGAAATGGCACCTGAAC 1341

Qy    331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390

Db   1342 GACGTCATGCTGATGGGTGACTTCAACGCTGACTGCTCTTATGTAACCTCTTCTCAGTGG 1401

Qy    391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450

Db   1402 TCATCGATTCGTCTGCGCACCTCGTCGACCTTCCAGTGGCTGATCCCGGACTCCGCTGAC 1461

Qy    451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510

Db   1462 ACCACCGCTACT---AGTACCAACTGCGCTTACGACCGTATCGTTGTTGCTGGATCCCTG 1518

Qy    511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570

Db   1519 CTGCAGTCTTCTGTTGTACCGGGTAGCGCGGCCCCGTTCGACTTCCAGGCTGCATATGGT 1578

Qy    571 TTGTCTGAATCGAAGGCCCTGG 592

Db   1579 CTTTCGAACGAAATGGCGCTGG 1600


RESULT 11
US-08-468-012A-1
; Sequence 1, Application US/08468012A
; Patent No. 5830744
;  GENERAL INFORMATION:
;    APPLICANT:  RUBEN, ET AL.
;    TITLE OF INVENTION:  Human DNase
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,012A
;      FILING DATE:  6 JUNE 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04954
```

; FILING DATE: 5 MAY 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-428
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1055 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: cDNA
US-08-468-012A-1

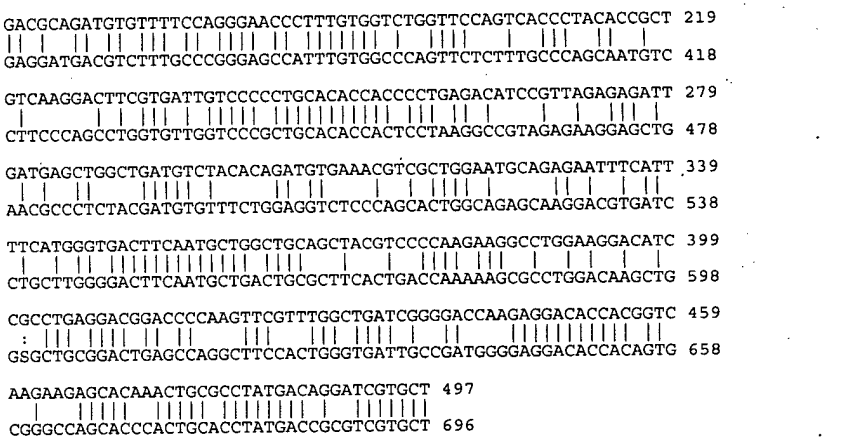
Query Match 19.5%; Score 115.2; DB 2; Length 1055;
Best Local Similarity 58.6%; Pred. No. 2.2e-26;
Matches 198; Conservative 1; Mismatches 139; Indels 0; Gaps 0;



RESULT 12
US-09-054-989-1
; Sequence 1, Application US/09054989
; Patent No. 6251648
; GENERAL INFORMATION:
; APPLICANT: RUBEN, ET AL.
; TITLE OF INVENTION: Human DNase
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/054,989
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/468,012

; FILING DATE: 6-JUNE-1995
; APPLICATION NUMBER: PCT/US94/04954
; FILING DATE: 5 MAY 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-428
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1055 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: cDNA
US-09-054-989-1

Query Match 19.5%; Score 115.2; DB 4; Length 1055;
Best Local Similarity 58.6%; Pred. No. 2.2e-26;
Matches 198; Conservative 1; Mismatches 139; Indels 0; Gaps 0;



RESULT 13
US-09-173-581-10
; Sequence 10, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/173,581A
; CURRENT FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -

```
;   OTHER INFORMATION: 119819
US-09-173-581-10

  Query Match             5.9%;  Score 35.2;  DB 3;  Length 1427;
  Best Local Similarity   53.7%;  Pred. No. 0.28;
  Matches   73;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy     322 AATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAG 381
           ||||||||| ||  |||   ||| |   |     |   |     | | |||  || | ||||
Db     857 AATGCAGATAAGATCAAGATCAAGATCGCAGACCTGGGCAACGCCTGCTGGGTGCACAAG 916

Qy     382 AAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGAC 441
           |    |  | | |||||||      ||  | |||       |||    ||||||||| | |
Db     917 CACTTCACGGAAGACATCCAGACTCGGCAGTACCGGGCCGTCGAGGTGCTGATCGGCGCC 976

Qy     442 CAAGAGGACACCACGG 457
           || | | | || |||
Db     977 GAATACGGCCCCCCGG 992


RESULT 14
US-09-420-915-10
; Sequence 10, Application US/09420915
; Patent No. 6264947
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Yue, Henry
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Gorgone, Gina
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Lu, Aina
;  TITLE OF INVENTION: Protein Kinase Homologs
;  FILE REFERENCE: PF-0614 US
;  CURRENT APPLICATION NUMBER: US/09/420,915
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/173,581
;  EARLIER FILING DATE: 1998-10-15
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PERL Program
; SEQ ID NO 10
;   LENGTH: 1427
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: 119819
US-09-420-915-10

  Query Match             5.9%;  Score 35.2;  DB 4;  Length 1427;
  Best Local Similarity   53.7%;  Pred. No. 0.28;
  Matches   73;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy     322 AATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAG 381
           ||||||||| ||  |||   ||| |   |     |   |     | | |||  || | ||||
Db     857 AATGCAGATAAGATCAAGATCAAGATCGCAGACCTGGGCAACGCCTGCTGGGTGCACAAG 916

Qy     382 AAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGAC 441
           |    |  | | |||||||      ||  | |||       |||    ||||||||| | |
Db     917 CACTTCACGGAAGACATCCAGACTCGGCAGTACCGGGCCGTCGAGGTGCTGATCGGCGCC 976

Qy     442 CAAGAGGACACCACGG 457
           || | | | || |||
Db     977 GAATACGGCCCCCCGG 992


RESULT 15
US-08-924-440-1
; Sequence 1, Application US/08924440
; Patent No. 5871550
;  GENERAL INFORMATION:
;    APPLICANT:  Frits et al.
;    TITLE OF INVENTION:  MUTANT THERMONOSPORA SPP. CELLULASE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International, Inc.
;      STREET:  925 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/924,440
;      FILING DATE:  August 27, 1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Christopher L. Stone
;      REGISTRATION NUMBER:  35,696
;      REFERENCE/DOCKET NUMBER:  GC388
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 846-7555
;      TELEFAX:  (650) 845-6504
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1293 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-924-440-1

  Query Match             5.5%;  Score 32.8;  DB 2;  Length 1293;
  Best Local Similarity   58.0%;  Pred. No. 1.5;
  Matches   58;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy     351 CTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGAC 410
           .|| ||  | || | ||  | ||| |  | ||| ||||       |||||  ||||| | |
Db    1080 CTACACCGGTGACGGCGCCAACGACTTCCAGATGGCCGACCGCTACATCGACCTGATGGC 1139

Qy     411 GGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450
           ||| |  ||| |||  |||    |   || || |   ||||
Db    1140 GGAACGGAAGATCGGGTGGACCAAGTGGAACTACTCGGAC 1179


Search completed: December  9, 2002, 09:19:41
Job time : 52.0918 secs
```

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:00:58 ; Search time 117.65 Seconds
(without alignments)
1962.092 Million cell updates/sec

Title: US-09-905-114-1
Perfect score: 592
Sequence: 1 gagaagctaaacggaaattc.....gtctgaatcgaaggccctgg 592

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	592	100.0	592	10	US-09-905-114-1	Sequence 1, Appli
2	592	100.0	671	10	US-09-905-114-3	Sequence 3, Appli
3	476.8	80.5	1023	10	US-09-962-436-336	Sequence 336, App
4	476.8	80.5	1023	10	US-09-954-456-271	Sequence 271, App
5	476.8	80.5	1023	10	US-09-880-107-3392	Sequence 3392, Ap
6	476.8	80.5	1079	12	US-10-133-065-1	Sequence 1, Appli
7	476.8	80.5	1079	12	US-10-074-509-1	Sequence 1, Appli
8	435.2	73.5	1124	12	US-10-133-065-11	Sequence 11, Appl
9	435.2	73.5	1124	12	US-10-074-509-11	Sequence 11, Appl
c 10	171.4	29.0	392	10	US-09-880-107-630	Sequence 630, App
11	149.4	25.2	1039	10	US-09-771-078-1	Sequence 1, Appli
12	149.4	25.2	1039	10	US-09-825-012-2	Sequence 2, Appli
13	147.8	25.0	783	10	US-09-825-012-4	Sequence 4, Appli
14	147.8	25.0	858	10	US-09-825-012-6	Sequence 6, Appli
15	147.8	25.0	1548	10	US-09-825-012-62	Sequence 62, Appl
16	147.8	25.0	1548	10	US-09-825-012-63	Sequence 63, Appl
17	147.8	25.0	1548	10	US-09-825-012-65	Sequence 65, Appl
18	147.8	25.0	1554	10	US-09-825-012-34	Sequence 34, Appl
19	147.8	25.0	1554	10	US-09-825-012-35	Sequence 35, Appl
20	147.8	25.0	1554	10	US-09-825-012-37	Sequence 37, Appl
21	147.8	25.0	1557	10	US-09-825-012-64	Sequence 64, Appl
22	147.8	25.0	1560	10	US-09-825-012-72	Sequence 72, Appl
23	147.8	25.0	1560	10	US-09-825-012-73	Sequence 73, Appl
24	147.8	25.0	1560	10	US-09-825-012-75	Sequence 75, Appl
25	147.8	25.0	1560	10	US-09-825-012-77	Sequence 77, Appl
26	147.8	25.0	1560	10	US-09-825-012-78	Sequence 78, Appl
27	147.8	25.0	1560	10	US-09-825-012-79	Sequence 79, Appl
28	147.8	25.0	1563	10	US-09-825-012-36	Sequence 36, Appl
29	147.8	25.0	1566	10	US-09-825-012-67	Sequence 67, Appl
30	147.8	25.0	1566	10	US-09-825-012-68	Sequence 68, Appl
31	147.8	25.0	1566	10	US-09-825-012-70	Sequence 70, Appl
32	147.8	25.0	1569	10	US-09-825-012-74	Sequence 74, Appl
33	147.8	25.0	1575	10	US-09-825-012-69	Sequence 69, Appl
34	147.8	25.0	1578	10	US-09-825-012-81	Sequence 81, Appl
35	147.8	25.0	1578	10	US-09-825-012-82	Sequence 82, Appl
36	147.8	25.0	1578	10	US-09-825-012-84	Sequence 84, Appl
37	147.8	25.0	1584	10	US-09-825-012-39	Sequence 39, Appl
38	147.8	25.0	1584	10	US-09-825-012-40	Sequence 40, Appl
39	147.8	25.0	1584	10	US-09-825-012-42	Sequence 42, Appl
40	147.8	25.0	1587	10	US-09-825-012-83	Sequence 83, Appl
41	147.8	25.0	1590	10	US-09-825-012-91	Sequence 91, Appl
42	147.8	25.0	1590	10	US-09-825-012-92	Sequence 92, Appl
43	147.8	25.0	1590	10	US-09-825-012-94	Sequence 94, Appl
44	147.8	25.0	1593	10	US-09-825-012-41	Sequence 41, Appl
45	147.8	25.0	1596	10	US-09-825-012-86	Sequence 86, Appl

ALIGNMENTS

```
RESULT 1
US-09-905-114-1
; Sequence 1, Application US/09905114
; Patent No. US20020048745A1
; GENERAL INFORMATION:
;  APPLICANT: ZHANG, HUANMIN
;  APPLICANT:  AX, ROY L
;  APPLICANT:  BELLIN, MARY E
;  TITLE OF INVENTION: ISOLATED POLYNUCLEOTIDE SEQUENCES ENCODING A FERTILITY ASSOCIATED
;  FILE REFERENCE: 210707US20
;  CURRENT APPLICATION NUMBER: US/09/905,114
;  CURRENT FILING DATE:  2001-07-14
;  PRIOR APPLICATION NUMBER: US 60/218,140
;  PRIOR FILING DATE: 2000-07-14
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Bos sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(591)
;   OTHER INFORMATION:
US-09-905-114-1

  Query Match           100.0%;  Score 592;  DB 10;  Length 592;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-176;
  Matches  592;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
```

```
Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
```

RESULT 2
US-09-905-114-3
; Sequence 3, Application US/09905114
; Patent No. US20020048745A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, HUANMIN
; APPLICANT: AX, ROY L
; APPLICANT: BELLIN, MARY E
; TITLE OF INVENTION: ISOLATED POLYNUCLEOTIDE SEQUENCES ENCODING A FERTILITY ASSOCIATED
; FILE REFERENCE: 210707US20
; CURRENT APPLICATION NUMBER: US/09/905,114
; CURRENT FILING DATE: 2001-07-14
; PRIOR APPLICATION NUMBER: US 60/218,140
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 671
; TYPE: DNA
; ORGANISM: Bos sp.
US-09-905-114-3

Query Match 100.0%; Score 592; DB 10; Length 671;
Best Local Similarity 100.0%; Pred. No. 1.3e-176;
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     27 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 86

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 146

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    147 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 206

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 266

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    267 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 326

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 386

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 446

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 506

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    507 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 566

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    567 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 618
```

RESULT 3
US-09-962-436-336
; Sequence 336, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 336
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-336

Query Match 80.5%; Score 476.8; DB 10; Length 1023;
Best Local Similarity 87.8%; Pred. No. 3.1e-140;
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||| |||||||||||||||||||||||||
Db    241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db    301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||||
Db    361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db    421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          |||||||||||||||| |||||||||||  ||||| |||||| ||| ||| ||||||
Db    481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||||
Db    541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600
```

```
Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
           |||||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db     601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
           || ||||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db     661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
           |||||||||||| |||||||||||||||| | || ||||  |||| || |||  | ||||||
Db     721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
             || ||||| |||||||||||||||| | || |||||  | |||||||||
Db     781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832
```

RESULT 4
US-09-954-456-271
; Sequence 271, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 271
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-271

Query Match 80.5%; Score 476.8; DB 10; Length 1023;
Best Local Similarity 87.8%; Pred. No. 3.1e-140;
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
           |||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||||
Db     241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
           |||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db     301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
           || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||||
Db     361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
           || |||||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
           |||||||||||||||| ||||||||||||  ||||| |||||| ||| ||| ||||||
Db     481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
           || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||||
Db     541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
           |||||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db     601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
           || ||||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db     661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
           |||||||||||| |||||||||||||||| | || ||||  |||| || |||  | ||||||
Db     721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
             || ||||| |||||||||||||||| | || |||||  | |||||||||
Db     781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832
```

RESULT 5
US-09-880-107-3392
; Sequence 3392, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3392
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U56814
US-09-880-107-3392

Query Match 80.5%; Score 476.8; DB 10; Length 1023;
Best Local Similarity 87.8%; Pred. No. 3.1e-140;
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
           |||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||||
Db     241 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 300

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
           |||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db     301 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 360

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
           || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||||
Db     361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 420

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
```

```
          || |||||||||||||||||||| || ||| |||| |||||||| ||||||||||| ||
Db    421 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 480

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||||||||||||||||| ||||||||||||  ||||| |||||| ||| ||| ||||||
Db    481 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 540

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db    541 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 600

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db    601 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 660

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db    661 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 720

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||| | || ||||  |||| || |||  | ||||||
Db    721 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            || ||||| |||||||||||||||| | || |||||  | |||||||||
Db    781 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 832
```

RESULT 6
US-10-133-065-1
; Sequence 1, Application US/10133065
; Patent No. US20020123122A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; Baron, Will F.
; TITLE OF INVENTION: HUMAN DNASE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/133,065
; FILING DATE: 26-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/643,520
; FILING DATE: 22/08/2000
; APPLICATION NUMBER: 08/794827
; FILING DATE: 04-Feb-1997
; APPLICATION NUMBER: 60/109796
; FILING DATE: 05-Feb-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Evans, David W
; REGISTRATION NUMBER: *See attached Limited Recognition under 37
; C.F.R. 10.9(b)
; REFERENCE/DOCKET NUMBER: P1000R1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1739
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1079 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-133-065-1

Query Match 80.5%; Score 476.8; DB 12; Length 1079;
Best Local Similarity 87.8%; Pred. No. 3.2e-140;
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||| |||||||||||||||||||||||
Db    287 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 346

Qy     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db    347 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 406

Qy    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||     || || | |||||| |||||||||| ||||||||||||||||||||||||
Db    407 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 466

Qy    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db    467 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 526

Qy    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          |||||||||||||||| ||||||||||||  ||||| |||||| ||| ||| ||||||
Db    527 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 586

Qy    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db    587 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 646

Qy    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db    647 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 706

Qy    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || |||||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db    707 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 766

Qy    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||| | || ||||  |||| || |||  | ||||||
Db    767 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 826

Qy    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
            || ||||| |||||||||||||||| | || |||||  | |||||||||
Db    827 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 878
```

RESULT 7
US-10-074-509-1
; Sequence 1, Application US/10074509
; Patent No. US20020142437A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; Baron, Will F.
; TITLE OF INVENTION: HUMAN DNASE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/074,509
; FILING DATE: 11-Feb-2002

```
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/643,520
;            FILING DATE: 22/08/2000
;            APPLICATION NUMBER: 08/794827
;            FILING DATE: 04-Feb-1997
;            APPLICATION NUMBER: 60/109796
;            FILING DATE: 05-Feb-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Johnston, Sean A.
;            REGISTRATION NUMBER: 35,910
;            REFERENCE/DOCKET NUMBER: P1000R1C2
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650/225-3562
;            TELEFAX: 650/952-9881
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1079 base pairs
;            TYPE: Nucleic Acid
;            STRANDEDNESS: Single
;            TOPOLOGY: Linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-074-509-1

  Query Match             80.5%;  Score 476.8;  DB 12;  Length 1079;
  Best Local Similarity   87.8%;  Pred. No. 3.2e-140;
  Matches  520;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| ||| |||||||||| | ||||||||| ||||||||||||||||||||||||
Db     287 GAGAAGCTGAACAGAAATTCAAGGAGAGGCATAACGTACAACTATGTGATTAGCTCTCGG 346

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          ||||||||||||||||||||||||| |||||||||||||| || |||||||| ||||||
Db     347 CTTGGAAGAAACACATATAAAGAACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCT 406

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||    || || | |||||| |||||||||| ||||||||||||||||||||||||
Db     407 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGG 466

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||||||||||| || ||| |||| |||||||| |||||||||||| ||
Db     467 GAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTATC 526

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          |||||||||||||||| |||||||||||||  ||||| |||||| ||| ||| ||||||
Db     527 CCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTAC 586

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || || |||||||  |||||||| || ||||||||||||||||||||||||||||||||||
Db     587 ACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCC 646

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
          |||||||||||||||||||||||||||||||| |||||||| ||||||| ||||||| |
Db     647 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGG 706

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
          || ||||||||||||||||||||||||||||||||| |||||||||| ||||| ||
Db     707 TTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCA 766

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
          |||||||||| |||||||||||||| | || ||||  |||| || |||  | ||||||
Db     767 TATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 826

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
             || ||||| |||||||||||||||| | || |||||  | |||||||||
Db     827 AGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAGGCCCTGG 878


RESULT 8
US-10-133-065-11
; Sequence 11, Application US/10133065
; Patent No. US20020123122A1
;  GENERAL INFORMATION:
;       APPLICANT: Baker, Kevin P.
;                  Baron, Will F.
;       TITLE OF INVENTION: HUMAN DNASE
;       NUMBER OF SEQUENCES: 11
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genentech, Inc.
;            STREET: 1 DNA Way
;            CITY: South San Francisco
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94080
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: WinPatin (Genentech)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/133,065
;            FILING DATE: 26-Apr-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/643,520
;            FILING DATE: 22/08/2000
;            APPLICATION NUMBER: 08/794827
;            FILING DATE: 04-Feb-1997
;            APPLICATION NUMBER: 60/109796
;            FILING DATE: 05-Feb-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Evans, David W
;            REGISTRATION NUMBER: *See attached Limited Recognition under 37
;                                 C.F.R. 10.9(b)
;            REFERENCE/DOCKET NUMBER: P1000R1C3
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650/225-1739
;            TELEFAX: 650/952-9881
;  INFORMATION FOR SEQ ID NO: 11:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1124 base pairs
;            TYPE: Nucleic Acid
;            STRANDEDNESS: Single
;            TOPOLOGY: Linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-133-065-11

  Query Match             73.5%;  Score 435.2;  DB 12;  Length 1124;
  Best Local Similarity   83.4%;  Pred. No. 4.2e-127;
  Matches  494;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
          |||||||| || ||||||||| ||| | ||| ||||||||||||||||||||| |||||
Db     404 GAGAAGCTGAATGGAAATTCACGAAGAAGCACAACATACAACTATGTGATTAGTTCTCGA 463

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
          |||||||||||||| || ||||| |||||||||||  |||| || || ||||| ||||||
Db     464 CTTGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGCTGGTGTCT 523

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
          || ||   ||  |||| |||||| |||||||||| ||||||| |||| ||||||||||||
Db     524 GTGAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 583

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
          || |||||||||| |||||||| || ||||  || |||||||||||||||||||||||||
Db     584 GAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCGTGATTGTC 643

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
          ||| ||||||| || || ||||| ||||||| |||||| |||||||||  |||||||||
Db     644 CCCTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCGATGTCTAC 703

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
          || ||||||| | | |  |||||  |||||||||||||| |||||||||| ||||| ||
Db     704 ACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATTTCAACGCC 763
```

```
Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
           ||||| ||||| |||||||||||||||||| || |||| ||  ||||||||||||||||||
Db     764 GGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGAGGACGGACCCCAAG 823

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
           || |||||||||| |||||||||||||||| ||||||||||||| || | ||| |||
Db     824 TTTGTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCAGCTGTGCC 883

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
           |||||||||| |||||| | |||||| | || |||||||| |  | ||| |  || |
Db     884 TATGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTTCCCCGTTCCAGT 943

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
           |||||||||  || ||||||||||    ||||||||   | |||||||||
Db     944 GGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 995


RESULT 9
US-10-074-509-11
; Sequence 11, Application US/10074509
; Patent No. US20020142437A1
;  GENERAL INFORMATION:
;       APPLICANT: Baker, Kevin P.
;                  Baron, Will F.
;       TITLE OF INVENTION: HUMAN DNASE
;       NUMBER OF SEQUENCES: 11
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genentech, Inc.
;            STREET: 1 DNA Way
;            CITY: South San Francisco
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94080
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: WinPatin (Genentech)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/074,509
;            FILING DATE: 11-Feb-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/643,520
;            FILING DATE: 22/08/2000
;            APPLICATION NUMBER: 08/794827
;            FILING DATE: 04-Feb-1997
;            APPLICATION NUMBER: 60/109796
;            FILING DATE: 05-Feb-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Johnston, Sean A.
;            REGISTRATION NUMBER: 35,910
;            REFERENCE/DOCKET NUMBER: P1000R1C2
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650/225-3562
;            TELEFAX: 650/952-9881
;  INFORMATION FOR SEQ ID NO: 11:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1124 base pairs
;            TYPE: Nucleic Acid
;            STRANDEDNESS: Single
;            TOPOLOGY: Linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-074-509-11

  Query Match             73.5%;  Score 435.2;  DB 12;  Length 1124;
  Best Local Similarity   83.4%;  Pred. No. 4.2e-127;
  Matches  494;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy       1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60
           |||||||| || ||||||||| ||| | ||| ||||||||||||||||||||| |||||
Db     404 GAGAAGCTGAATGGAAATTCACGAAGAAGCACAACATACAACTATGTGATTAGTTCTCGA 463

Qy      61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120
           |||||||||||||| || ||||| |||||||||||  |||| || || ||||| ||||||
Db     464 CTTGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGCTGGTGTCT 523

Qy     121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180
           || ||   ||  |||| |||||| |||||||||| ||||||| |||| ||||||||||||
Db     524 GTGAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG 583

Qy     181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
           || |||||||||| |||||||| || ||||  || ||||||||||||||||||||||||||
Db     584 GAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCGTGATTGTC 643

Qy     241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300
           ||| ||||||| || || ||||| ||||||| |||||| |||||||||  |||||||||
Db     644 CCCTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCGATGTCTAC 703

Qy     301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360
           || ||||||| | | |  |||||  |||||||||||| |||||||||| ||||| ||
Db     704 ACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATTTCAACGCC 763

Qy     361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420
           ||||| ||||| |||||||||||||||||| || |||| ||  ||||||||||||||||||
Db     764 GGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGAGGACGGACCCCAAG 823

Qy     421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480
           || |||||||||| |||||||||||||||| ||||||||||||| || | ||| |||
Db     824 TTTGTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCAGCTGTGCC 883

Qy     481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540
           |||||||||| |||||| | |||||| | || |||||||| |  | ||| |  || |
Db     884 TATGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTTCCCCGTTCCAGT 943

Qy     541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
           |||||||||  || ||||||||||    ||||||||   | |||||||||
Db     944 GGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 995


RESULT 10
US-09-880-107-630/c
; Sequence 630, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 630
;   LENGTH: 392
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA279676
US-09-880-107-630

  Query Match             29.0%;  Score 171.4;  DB 10;  Length 392;
  Best Local Similarity   86.0%;  Pred. No. 2.8e-44;
  Matches  190;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy     372 CGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCT 431
           |||||||||||||||||||||| |||||||| ||||||| ||||||| ||| ||||||||
Db     392 CGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCT 333

Qy     432 GATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGAT 491
```

```
Db    332 GATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCATATGACAGGAT 273

Qy    492 CGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGA 551

Db    272 TGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGA 213

Qy    552 TTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592

Db    212 CTTCCAGAAAGCTTACAAGCTGACTGAAAAGGAGGCCCTGG 172


RESULT 11
US-09-771-078-1
; Sequence 1, Application US/09771078
; Patent No. US20020034727A1
;  GENERAL INFORMATION:
;    APPLICANT:  Mrsny, Randall J.,
;    APPLICANT:  Daugherty, Ann L.,
;    APPLICANT:  Patapoff, Thomas W.
;    TITLE OF INVENTION:  DNase Compaction Assay
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/771,078
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/971,019
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Glaister, Debra J.
;      REGISTRATION NUMBER:  33,888
;      REFERENCE/DOCKET NUMBER:  792
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3177
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1039 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-771-078-1

  Query Match             25.2%;  Score 149.4;  DB 10;  Length 1039;
  Best Local Similarity   56.7%;  Pred. No. 4.1e-37;
  Matches  317;  Conservative    0;  Mismatches  236;  Indels    6;  Gaps    2;

Qy     34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93

Db    409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy     94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 150

Db    469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy    151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210

Db    529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy    211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270

Db    589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy    271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330

Db    649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy    331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390

Db    709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy    391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450

Db    769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy    451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510

Db    829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy    511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570

Db    886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy    571 TTGTCTGAATCGAAGGCCC 589

Db    946 CTGAGTGACCAACTGGCCC 964


RESULT 12
US-09-825-012-2
; Sequence 2, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Robert
;  TITLE OF INVENTION: Compounds for Targeting
;  FILE REFERENCE: 43191-256808
;  CURRENT APPLICATION NUMBER: US/09/825,012
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/237,159
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: GB 0008049.9
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1039
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-012-2

  Query Match             25.2%;  Score 149.4;  DB 10;  Length 1039;
  Best Local Similarity   56.7%;  Pred. No. 4.1e-37;
  Matches  317;  Conservative    0;  Mismatches  236;  Indels    6;  Gaps    2;

Qy     34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93

Db    409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy     94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 150

Db    469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy    151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210

Db    529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy    211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270

Db    589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy    271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330
```

```
Db    649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy    331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390

Db    709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy    391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450

Db    769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy    451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510

Db    829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy    511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570

Db    886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy    571 TTGTCTGAATCGAAGGCCC 589

Db    946 CTGAGTGACCAACTGGCCC 964


RESULT 13
US-09-825-012-4
; Sequence 4, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Robert
;  TITLE OF INVENTION: Compounds for Targeting
;  FILE REFERENCE: 43191-256808
;  CURRENT APPLICATION NUMBER: US/09/825,012
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/237,159
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: GB 0008049.9
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 783
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-012-4

  Query Match           25.0%;  Score 147.8;  DB 10;  Length 783;
  Best Local Similarity   56.5%;  Pred. No. 1.1e-36;
  Matches  316;  Conservative    0;  Mismatches  237;  Indels    6;  Gaps    2;

Qy     34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93

Db    184 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 243

Qy     94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 150

Db    244 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 303

Qy    151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210

Db    304 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 363

Qy    211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270

Db    364 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 423

Qy    271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330

Db    424 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 483

Qy    331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390

Db    484 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 543

Qy    391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450

Db    544 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 603

Qy    451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510

Db    604 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 660

Qy    511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570

Db    661 CTCCGAGGGGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 720

Qy    571 TTGTCTGAATCGAAGGCCC 589

Db    721 CTGAGTGACCAACTGGCCC 739


RESULT 14
US-09-825-012-6
; Sequence 6, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Robert
;  TITLE OF INVENTION: Compounds for Targeting
;  FILE REFERENCE: 43191-256808
;  CURRENT APPLICATION NUMBER: US/09/825,012
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/237,159
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: GB 0008049.9
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-012-6

  Query Match           25.0%;  Score 147.8;  DB 10;  Length 858;
  Best Local Similarity   56.5%;  Pred. No. 1.2e-36;
  Matches  316;  Conservative    0;  Mismatches  237;  Indels    6;  Gaps    2;

Qy     34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93

Db    259 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 318

Qy     94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 150

Db    319 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 378

Qy    151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210

Db    379 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 438

Qy    211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270

Db    439 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 498

Qy    271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330

Db    499 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 558

Qy    331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390

Db    559 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 618

Qy    391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450

Db    619 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 678

Qy    451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510
```

```
Db    679 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 735

Qy    511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570

Db    736 CTCCGAGGGGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 795

Qy    571 TTGTCTGAATCGAAGGCCC 589

Db    796 CTGAGTGACCAACTGGCCC 814


RESULT 15
US-09-825-012-62
; Sequence 62, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Robert
;  TITLE OF INVENTION: Compounds for Targeting
;  FILE REFERENCE: 43191-256808
;  CURRENT APPLICATION NUMBER: US/09/825,012
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/237,159
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: GB 0008049.9
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
;   LENGTH: 1548
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Humanised HMFG1 Fd - DNase I fusion
US-09-825-012-62

  Query Match           25.0%;  Score 147.8;  DB 10;  Length 1548;
  Best Local Similarity  56.5%;  Pred. No. 1.7e-36;
  Matches  316;  Conservative    0;  Mismatches  237;  Indels    6;  Gaps    2;

Qy     34 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 93

Db    949 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 1008

Qy     94 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 150

Db   1009 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 1068

Qy    151 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 210

Db   1069 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 1128

Qy    211 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 270

Db   1129 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 1188

Qy    271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 330

Db   1189 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 1248

Qy    331 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 390

Db   1249 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 1308

Qy    391 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 450

Db   1309 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 1368

Qy    451 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 510

Db   1369 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 1425

Qy    511 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 570

Db   1426 CTCCGAGGGGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 1485

Qy    571 TTGTCTGAATCGAAGGCCC 589

Db   1486 CTGAGTGACCAACTGGCCC 1504


Search completed: December  9, 2002, 09:24:07
Job time : 127.65 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:18:16 ; Search time 1408.41 Seconds
(without alignments)
13865.275 Million cell updates/sec

Title: US-09-905-114-3
Perfect score: 671
Sequence: 1 acaacaggatctgccccata.....atcatcatgaagaaccatga 671

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	527	78.5	1023	6	AX329742	AX329742 Sequence
2	527	78.5	1023	6	AX332286	AX332286 Sequence
3	527	78.5	1023	6	AX332789	AX332789 Sequence
4	527	78.5	1023	6	AX410746	AX410746 Sequence
5	527	78.5	1023	9	HSU56814	U56814 Human DNase
6	527	78.5	1067	9	BC015831	BC015831 Homo sapi
7	527	78.5	1079	9	AF047354	AF047354 Homo sapi
8	525.4	78.3	1108	9	HSU75744	U75744 Homo sapien
9	501.4	74.7	1331	10	AF039852	AF039852 Rattus no
10	501.4	74.7	1417	10	RNU75689	U75689 Rattus norv
11	481.2	71.7	1208	6	AR047845	AR047845 Sequence
12	481.2	71.7	1208	6	E11687	E11687 cDNA encodi
13	474.8	70.8	941	10	MMU76110	U76110 Mus musculu
14	474.8	70.8	1124	10	AF047355	AF047355 Mus muscu
15	473.2	70.5	2110	10	BC012671	BC012671 Mus muscu
16	411.4	61.3	1223	6	AX400016	AX400016 Sequence
17	298.4	44.5	1250	5	AF059612	AF059612 Xenopus l
c 18	195.6	29.2	392	6	AX330074	AX330074 Sequence
c 19	195.6	29.2	392	6	AX407983	AX407983 Sequence
20	182.4	27.2	1295	4	BTPANDNAI	AJ001538 Bos tauru
21	181.4	27.0	868	4	AF311922	AF311922 Bos tauru
22	173.8	25.9	1283	5	AB041732	AB041732 Cynops py
23	172.6	25.7	1210	5	AB038776	AB038776 Rana cate
24	164.8	24.6	1161	4	D82875	D82875 Oryctolagus
25	163.2	24.3	1039	6	AX030575	AX030575 Sequence
c 26	163.2	24.3	1039	6	AX030591	AX030591 Sequence
27	163.2	24.3	1039	6	AX268671	AX268671 Sequence
28	163.2	24.3	1039	9	HUMDNASEI	M55983 Human DNase
29	163.2	24.3	1149	9	BC029437	BC029437 Homo sapi
30	163.2	24.3	2054	9	HSA298844	AJ298844 Homo sapi
31	161.6	24.1	783	6	AX268673	AX268673 Sequence
32	161.6	24.1	858	6	AX268675	AX268675 Sequence
33	161.6	24.1	1039	6	AR019468	AR019468 Sequence
34	161.6	24.1	1312	10	BC014718	BC014718 Mus muscu
35	161.6	24.1	1337	10	BC030394	BC030394 Mus muscu
36	161.6	24.1	1548	6	AX268721	AX268721 Sequence
37	161.6	24.1	1548	6	AX268722	AX268722 Sequence
38	161.6	24.1	1554	6	AX268701	AX268701 Sequence
39	161.6	24.1	1554	6	AX268702	AX268702 Sequence
40	161.6	24.1	1557	6	AX268723	AX268723 Sequence
41	161.6	24.1	1560	6	AX268729	AX268729 Sequence
42	161.6	24.1	1560	6	AX268730	AX268730 Sequence
43	161.6	24.1	1560	6	AX268733	AX268733 Sequence
44	161.6	24.1	1560	6	AX268734	AX268734 Sequence
45	161.6	24.1	1563	6	AX268703	AX268703 Sequence

ALIGNMENTS

RESULT 1
AX329742
LOCUS AX329742 1023 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 251 from Patent WO0194629.
ACCESSION AX329742
VERSION AX329742.1 GI:18102720
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature

gene sets
JOURNAL Patent: WO 0194629-A 251 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1. .1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 312 a 244 c 240 g 227 t
ORIGIN

Query Match 78.5%; Score 527; DB 6; Length 1023;
Best Local Similarity 87.8%; Pred. No. 1.1e-132;
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db    275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db    335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||| |||||||||||||||||||| || ||| |||
Db    395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||| ||||||||||||  |
Db    455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db    515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| |||||||||||||||||||||||||||||||
Db    635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||||| ||||| || |||||||||||| |||||||||||||||| | || |
Db    695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db    755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||| |   | | |  ||
Db    815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869
```

RESULT 2
AX332286
LOCUS AX332286 1023 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 2795 from Patent WO0194629.
ACCESSION AX332286
VERSION AX332286.1 GI:18122920
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 2795 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1. .1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 312 a 244 c 240 g 227 t
ORIGIN

Query Match 78.5%; Score 527; DB 6; Length 1023;
Best Local Similarity 87.8%; Pred. No. 1.1e-132;
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db    275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db    335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||| |||||||||||||||||||| || ||| |||
Db    395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||| ||||||||||||  |
Db    455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db    515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| |||||||||||||||||||||||||||||||
Db    635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||||| ||||| || |||||||||||| |||||||||||||||| | || |
Db    695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db    755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||| |   | | |  ||
Db    815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869
```

RESULT 3
AX332789
LOCUS AX332789 1023 bp DNA linear PAT 10-JAN-2002
DEFINITION Sequence 3298 from Patent WO0194629.
ACCESSION AX332789
VERSION AX332789.1 GI:18123423
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,

```
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 3298 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .1023
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      312 a    244 c    240 g    227 t
ORIGIN

  Query Match             78.5%;  Score 527;  DB 6;  Length 1023;
  Best Local Similarity   87.8%;  Pred. No. 1.1e-132;
  Matches  575;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          |||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db     215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db     275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db     335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db     395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||| ||||||||||||  |
Db     455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db     515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||||
Db     635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||| ||||| || |||||||||| ||||||||||||||| | || |
Db     695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db     755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||||| |   | | |  ||
Db     815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869


RESULT 4
AX410746
LOCUS       AX410746               1023 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 3393 from Patent WO0229103.
ACCESSION   AX410746
VERSION     AX410746.1  GI:21443451
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 3393 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1. .1023
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. U56814"
BASE COUNT      312 a    244 c    240 g    227 t
ORIGIN

  Query Match             78.5%;  Score 527;  DB 6;  Length 1023;
  Best Local Similarity   87.8%;  Pred. No. 1.1e-132;
  Matches  575;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          |||||||||||||||||||||||||||||||||.|| ||| |||||||||| | ||||||||
Db     215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db     275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db     335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db     395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||| ||||||||||||  |
Db     455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db     515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||||
Db     635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||| ||||| || |||||||||| ||||||||||||||| | || |
Db     695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db     755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||||| |   | | |  ||
Db     815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869


RESULT 5
HSU56814
LOCUS       HSU56814               1023 bp    mRNA    linear   PRI 24-JUL-1997
DEFINITION  Human DNase1-Like III protein (DNAS1L3) mRNA, complete cds.
ACCESSION   U56814
VERSION     U56814.1  GI:1399718
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1023)
```

```
  AUTHORS   Rodriguez,A.M., Rodin,D., Nomura,H., Morton,C.C., Weremowicz,S. and
            Schneider,M.C.
  TITLE     Identification, localization, and expression of two novel human
            genes similar to deoxyribonuclease I
  JOURNAL   Genomics 42 (3), 507-513 (1997)
  MEDLINE   97349121
   PUBMED   9205125
REFERENCE   2  (bases 1 to 1023)
  AUTHORS   Schneider,M.C. and Rodriguez,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-1996) M.C. Schneider, Renal Division, Brigham and
            Women's Hospital, 75 Francis Street, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1. .1023
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Stratagene liver (#937224)"
                     /note="corresponds to EST clones 82269 (GenBank Accession
                     Number T68985, T69063), 82738 (GenBank Accession Number
                     T73558, T73653), 78422 (GenBank Accession Number T61400,
                     T61368)"
     gene            1. .1023
                     /gene="DNAS1L3"
     CDS             25. .942
                     /gene="DNAS1L3"
                     /codon_start=1
                     /product="DNase1-Like III protein"
                     /protein_id="AAB63967.1"
                     /db_xref="GI:1399719"
                     /translation="MSRELAPLLLLLLSIHSALAMRICSFNVRSFGESKQEDKNAMDV
                     IVKVIKRCDIILVMEIKDSNNRICPILMEKLNRNSRRGITYNYVISSRLGRNTYKEQY
                     AFLYKEKLVSVKRSYHYHDYQDGDADVFSREPFVVWFQSPHTAVKDFVIIPLHTTPET
                     SVKEIDELVEVYTDVKHRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPRFVWLIG
                     DQEDTTVKKSTNCAYDRIVLRGQEIVSSVVPKSNSVFDFQKAYKLTEEEALDVSDHFP
                     VEFKLQSSRAFTNSKKSVTLRKKTKSKRS"
BASE COUNT      312 a    244 c    240 g    227 t
ORIGIN

  Query Match             78.5%;  Score 527;  DB 9;  Length 1023;
  Best Local Similarity   87.8%;  Pred. No. 1.1e-132;
  Matches  575;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          |||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db     215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||| |||||||||||||||||||||||||| |||||||
Db     275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
           ||||||| || |||||||| |||||||| ||    || || | |||||| |||||||||
Db     335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
           |||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db     395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
           | |||||||| |||||||||| |||||||||||||||||| |||||||||||| |
Db     455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
           |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db     515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
           |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||
Db     635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
           |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db     695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
           ||| |||| || ||| | ||||||   || ||||| |||||||||||||||| | ||
Db     755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
           |||||  | |||||||||||||||||||||||||||||||| |   | | |  ||
Db     815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869


RESULT 6
BC015831
LOCUS       BC015831                1067 bp    mRNA    linear   PRI 29-OCT-2001
DEFINITION  Homo sapiens, deoxyribonuclease I-like 3, clone MGC:27146
            IMAGE:4723725, mRNA, complete cds.
ACCESSION   BC015831
VERSION     BC015831.1  GI:16198370
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1067)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
            Greene, Mark Ketteman and Anuradha Madan

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 38 Row: p Column: 7
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 4826697.
FEATURES             Location/Qualifiers
     source          1. .1067
                     /organism="Homo sapiens"
                     /db_xref="LocusID:1776"
                     /db_xref="taxon:9606"
                     /clone="MGC:27146 IMAGE:4723725"
                     /tissue_type="Liver"
                     /clone_lib="NIH_MGC_76"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     CDS             68. .985
                     /codon_start=1
                     /product="deoxyribonuclease I-like 3"
                     /protein_id="AAH15831.1"
                     /db_xref="GI:16198371"
                     /translation="MSRELAPLLLLLLSIHSALAMRICSFNVRSFGESKQEDKNAMDV
                     IVKVIKRCDIILVMEIKDSNNRICPILMEKLNRNSRRGITYNYVISSRLGRNTYKEQY
                     AFLYKEKLVSVKRSYHYHDYQDGDADVFSREPFVVWFQSPHTAVKDFVIIPLHTTPET
                     SVKEIDELVEVYTDVKHRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPRFVWLIG
                     DQEDTTVKKSTNCAYDRIVLRGQEIVSSVVPKSNSVFDFQKAYKLTEEEALDVSDHFP
                     VEFKLQSSRAFTNSKKSVTLRKKTKSKRS"
BASE COUNT      327 a    256 c    249 g    235 t
ORIGIN
```

```
  Query Match           78.5%;  Score 527;  DB 9;  Length 1067;
  Best Local Similarity  87.8%;  Pred. No. 1.1e-132;
  Matches  575;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    258 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 317

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||||| |||||||||||||||||||||||||||| |||||||
Db    318 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 377

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db    378 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 437

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db    438 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 497

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||| |||||||||||||||||| ||||||||||||   |
Db    498 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 557

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || ||||||| |||||||| || |||||||
Db    558 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 617

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    618 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 677

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||
Db    678 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 737

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db    738 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 797

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||  || ||||| |||||||||||||||| | ||
Db    798 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 857

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||||| |  | | |  ||
Db    858 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 912
```

```
RESULT 7
AF047354
LOCUS       AF047354               1079 bp    mRNA    linear   PRI 16-SEP-1998
DEFINITION  Homo sapiens liver and spleen DNase precursor (LSD) mRNA, complete
            cds.
ACCESSION   AF047354
VERSION     AF047354.1  GI:2905785
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Baron,W.F., Pan,C.Q., Spencer,S.A., Ryan,A.M., Lazarus,R.A. and
            Baker,K.P.
  TITLE     Cloning and characterization of an actin-resistant DNase I-like
            endonuclease secreted by macrophages
  JOURNAL   Gene 215 (2), 291-301 (1998)
  MEDLINE   98382522
   PUBMED   9714828
REFERENCE   2  (bases 1 to 1079)
  AUTHORS   Baron,W.F., Pan,C.Q., Spencer,S.A., Ryan,A.M., Lazarus,R.A. and
            Baker,K.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-1998) Molecular Biology, Genentech Inc., 1 DNA
            Way, South San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .1079
                     /gene="LSD"
     CDS             71. .988
                     /gene="LSD"
                     /note="actin resistant acivity; macrophage specific DNase
                     I; LS-DNase"
                     /codon_start=1
                     /product="liver and spleen DNase precursor"
                     /protein_id="AAC35752.1"
                     /db_xref="GI:2905786"
                     /translation="MSRELAPLLLLLLSIHSALAMRICSFNVRSFGESKQEDKNAMDV
                     IVKVIKRCDIILVMEIKDSNNRICPILMEKLNRNSRRGITYNYVISSRLGRNTYKEQY
                     AFLYKEKLVSVKRSYHYHDYQDGDADVFSREPFVVWFQSPHTAVKDFVIIPLHTTPET
                     SVKEIDELVEVYTDVKHRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPRFVWLIG
                     DQEDTTVKKSTNCAYDRIVLRGQEIVSSVVPKSNSVFDFQKAYKLTEEEALDVSDHFP
                     VEFKLQSSRAFTNSKKSVTLRKKTKSKRS"
     sig_peptide     71. .130
                     /gene="LSD"
                     /evidence=experimental
     mat_peptide     131. .985
                     /gene="LSD"
                     /product="liver and spleen DNase"
BASE COUNT      330 a    260 c    252 g    237 t
ORIGIN
```

```
  Query Match           78.5%;  Score 527;  DB 9;  Length 1079;
  Best Local Similarity  87.8%;  Pred. No. 1.1e-132;
  Matches  575;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    261 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 320

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||||| |||||||||||||||||||||||||||| |||||||
Db    321 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 380

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db    381 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 440

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db    441 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 500

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||| |||||||||||||||||| ||||||||||||   |
Db    501 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 560

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || ||||||| |||||||| || |||||||
Db    561 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 620

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    621 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 680

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||
Db    681 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 740

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db    741 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 800

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
```

```
Db    801 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 860

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
Db    861 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 915


RESULT 8
HSU75744
LOCUS       HSU75744                1108 bp    mRNA    linear   PRI 17-JUN-1998
DEFINITION  Homo sapiens DNase gamma mRNA, complete cds.
ACCESSION   U75744
VERSION     U75744.1  GI:3236319
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1108)
  AUTHORS   Shiokawa,D., Hirai,M. and Tanuma,S.
  TITLE     cDNA cloning of human DNase gamma: chromosomal localization of its
            gene and enzymatic properties of recombinant protein
  JOURNAL   Apoptosis 3 (2), 89-95 (1998)
REFERENCE   2  (bases 1 to 1108)
  AUTHORS   Shiokawa,D. and Tanuma,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-1996) Biochemistry, Science Univ. of Tokyo,
            Shinjuku-ku Ichigaya Funagawaracho, Tokyo 162, Japan
FEATURES             Location/Qualifiers
     source          1. .1108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             93. .1010
                     /function="apoptosis"
                     /function="deoxyribonuclease"
                     /codon_start=1
                     /product="DNase gamma"
                     /protein_id="AAC23652.1"
                     /db_xref="GI:3236320"
                     /translation="MSRELAPLLLLLLSIHSALAMRICSFNVRSFGESKQEDKNAMDV
                     IVKVIKRCDIILVMEIKDSNNRICPILMEKLNRNSRRGITYNYVISSRLGRKTYKEQY
                     AFLYKEKLVSVKRSYHYHDYQDGDADVFSREPFVVWFQSPHTAVKDFVIIPLHTTPET
                     SVKEIDELVEVYTDVKHRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPRFVWLIG
                     DQEDTTVKKSTNCAYDRIVLRGQEIVSSVVPKSNSVFDFQKAYKLTEEEALDVSDHFP
                     VEFKLQSSRAFTNSKKSVTLRKKTKSKRS"
BASE COUNT      329 a    261 c    271 g    247 t
ORIGIN

  Query Match             78.3%;  Score 525.4;  DB 9;  Length 1108;
  Best Local Similarity   87.6%;  Pred. No. 3.1e-132;
  Matches  574;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
Db    283 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 342

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
Db    343 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAAAACATATAAAGAACAATATGCCT 402

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
Db    403 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 462

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
Db    463 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 522

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
Db    523 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 582

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
Db    583 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 642

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
Db    643 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 702

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
Db    703 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 762

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
Db    763 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 822

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
Db    823 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 882

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
Db    883 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 937


RESULT 9
AF039852
LOCUS       AF039852                1331 bp    mRNA    linear   ROD 07-AUG-1998
DEFINITION  Rattus norvegicus DNaseY mRNA, complete cds.
ACCESSION   AF039852
VERSION     AF039852.1  GI:3395771
KEYWORDS    .
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1331)
  AUTHORS   Liu,Q.Y., Pandey,S., Singh,R.K., Lin,W., Ribecco,M.,
            Borowy-Borowski,H., Smith,B., LeBlanc,J., Walker,P.R. and
            Sikorska,M.
  TITLE     DNaseY: a rat DNaseI-like gene coding for a constitutively
            expressed chromatin-bound endonuclease
  JOURNAL   Biochemistry 37 (28), 10134-10143 (1998)
  MEDLINE   98332537
   PUBMED   9665719
REFERENCE   2  (bases 1 to 1331)
  AUTHORS   Liu,Q.Y., Singh,R.K., Lin,W. and Sikorska,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-DEC-1997) Institute for Biological Sciences, National
            Research Council, 1200 Montreal Road, Bldg. M-54, Box 4, Ottawa, ON
            K1A 0R6, Canada
FEATURES             Location/Qualifiers
     source          1. .1331
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /tissue_type="liver"
     CDS             153. .1085
                     /function="apoptotic endonuclease"
                     /note="similar to rat DNAse I; chromatin-bound"
                     /codon_start=1
                     /product="DNAseY"
                     /protein_id="AAC28937.1"
                     /db_xref="GI:3395772"
                     /translation="MSLYPASPYLASLLLFILALHGALSLRLCSFNVRSFGESKKENH
                     NAMDIIVKIIKRCDLILLMEIKDSNNNICPMLMEKLNGNSRRSTTYNYVISSRLGRNT
                     YKEQYAFLYKEKLVSVKAKYLYHDYQDGDTDVFSREPFVVWFQAPFTAAKDFVIVPLH
                     TTPETSVKEIDELADVYTDVRRRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPNF
                     VWLIGDQEDTTVKKSTSCAYDRIVLRGQEIVNSVVPRSSGVFDFQKAYELSEEEALDV
                     SDHFPVEFKLQSSRAFTNSRKSVSLKKKKKGSRS"
BASE COUNT      351 a    333 c    320 g    327 t
ORIGIN
```

Query Match 74.7%; Score 501.4; DB 10; Length 1331;
Best Local Similarity 85.3%; Pred. No. 1.1e-125;
Matches 559; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
Db     358 ACAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAACTCACGAAGAAGCACGA 417

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
Db     418 CATACAACTACGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAGTATGCCT 477

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
Db     478 TCCTCTACAAGGAGAAGCTGGTGTCTGTGAAGGCAAAATACCTCTACCATGACTATCAGG 537

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
Db     538 ATGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCA 597

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
Db     598 CTGCTGCCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCTGAAACCTCCGTTAAAG 657

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
Db     658 AGATAGATGAGCTGGCTGACGTCTACACGGATGTGAGAAGACGATGGAAGGCAGAGAATT 717

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
Db     718 TCATCTTCATGGGTGATTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 777

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
Db     778 ACATCCGTTTGAGGACAGACCCCAACTTTGTTTGGCTGATTGGGGACCAAGAGGACACCA 837

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
Db     838 CGGTCAAGAAGAGCACCAGCTGTGCCTATGACAGGATTGTGCTTCGCGGACAAGAGATAG 897

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
Db     898 TCAACTCTGTGGTTCCCCGCTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGT 957

Qy     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
Db     958 CTGAAGAGGAGGCCCTGGATGTCAGTGACCACTTTCCAGTTGAGTTTAAGCTACA 1012
```

RESULT 10
RNU75689
LOCUS RNU75689 1417 bp mRNA linear ROD 15-JUN-1998
DEFINITION Rattus norvegicus DNase gamma mRNA, complete cds.
ACCESSION U75689
VERSION U75689.1 GI:3220024
KEYWORDS .
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 1417)
AUTHORS Shiokawa,D., Ohyama,H., Yamada,T., Takahashi,K. and Tanuma,S.
TITLE Identification of an endonuclease responsible for apoptosis in rat thymocytes
JOURNAL Eur. J. Biochem. 226 (1), 23-30 (1994)
MEDLINE 95045594
PUBMED 7957253
REFERENCE 2 (bases 1 to 1417)
AUTHORS Shiokawa,D. and Tanuma,S.
TITLE Molecular cloning and expression of a cDNA encoding an apoptotic endonuclease DNase gamma
JOURNAL Biochem. J. 332 (Pt 3), 713-720 (1998)
MEDLINE 98285539
PUBMED 9620874
REFERENCE 3 (bases 1 to 1417)
AUTHORS Shiokawa,D. and Tanuma,S.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1996) Biochemistry, Science Univ. of Tokyo, Shinjuku-ku ichigaya funagawaracho, Tokyo 162, Japan
FEATURES Location/Qualifiers
source 1. .1417
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="spleen"
CDS 203. .1135
/function="apoptosis"
/note="deoxyribonuclease"
/codon_start=1
/product="DNase gamma"
/protein_id="AAC40134.1"
/db_xref="GI:3220025"
/translation="MSLYPASPYLASLLLFILALHGALSLRLCSFNVRSFGESKKENH
NAMDIIVKIIKRCDLILLMEIKDSNNNICPMLMEKLNGNSRRSTTYNYVISSRLGRNT
YKEQYAFLYKEKLVSVKAKYLYHDYQDGDTDVFSREPFVVWFQAPFTAAKDFVIVPLH
TTPETSVKEIDELADVYTDVRRRWKAENFIFMGDFNAGCSYVPKKAWKNIRLRTDPNF
VWLIGDQEDTTVKKSTSCAYDRIVLRGQEIVNSVVPRSSGVFDFQKAYELSEEEALDV
SDHFPVEFKLQSSRAFTNSRKSVSLKKKKKGSRS"
BASE COUNT 386 a 353 c 333 g 345 t
ORIGIN

Query Match 74.7%; Score 501.4; DB 10; Length 1417;
Best Local Similarity 85.3%; Pred. No. 1.1e-125;
Matches 559; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
Db     408 ACAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAACTCACGAAGAAGCACGA 467

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
Db     468 CATACAACTACGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAGTATGCCT 527

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
Db     528 TCCTCTACAAGGAGAAGCTGGTGTCTGTGAAGGCAAAATACCTCTACCATGACTATCAGG 587

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
Db     588 ATGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCA 647

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
Db     648 CTGCTGCCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCTGAAACCTCCGTTAAAG 707

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
Db     708 AGATAGATGAGCTGGCTGACGTCTACACGGATGTGAGAAGACGATGGAAGGCAGAGAATT 767

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
Db     768 TCATCTTCATGGGTGATTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 827

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
Db     828 ACATCCGTTTGAGGACAGACCCCAACTTTGTTTGGCTGATTGGGGACCAAGAGGACACCA 887

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
Db     888 CGGTCAAGAAGAGCACCAGCTGTGCCTATGACAGGATTGTGCTTCGCGGACAAGAGATAG 947

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
Db     948 TCAACTCTGTGGTTCCCCGCTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGT 1007
```

```
Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | ||||||||||||||| |||||||||||||| |   | | |  ||
Db   1008 CTGAAGAGGAGGCCCTGGATGTCAGTGACCACTTTCCAGTTGAGTTTAAGCTACA 1062


RESULT 11
AR047845
LOCUS       AR047845                1208 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 2 from patent US 5821103.
ACCESSION   AR047845
VERSION     AR047845.1  GI:5970188
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1208)
  AUTHORS   Tanuma,S.-i.
  TITLE     Deoxyribonuclease
  JOURNAL   Patent: US 5821103-A 2 13-OCT-1998;
FEATURES             Location/Qualifiers
     source          1. .1208
                     /organism="unknown"
BASE COUNT      319 a    304 c    283 g    302 t
ORIGIN

  Query Match             71.7%;  Score 481.2;  DB 6;  Length 1208;
  Best Local Similarity   85.1%;  Pred. No. 3.5e-120;
  Matches  550;  Conservative    0;  Mismatches   93;  Indels    3;  Gaps    1;

Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||  ||||| ||||| ||||||||||||| || ||||| ||| ||| | |||  |
Db    217 ACAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAACTCACGAAGAAGCACGA 276

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          |||||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    277 CATACAACTACGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAGTATGCCT 336

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          | ||||| || || ||||| |||||||| ||    ||  |||||||||| ||||||||||
Db    337 TCCTCTACAAGGAGAAGCTGGTGTCTGTGAAGGCAAAATACCTCTACCATGACTATCAGG 396

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||  |||| ||||||||||||| |||||||||| |||||||| | |||| ||
Db    397 ATGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCA 456

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||| |||||||||||||||||||||| ||||||| || ||||| || |||||||  ||
Db    457 CTGCTGCCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCTGAAACCTCCGTTAAAG 516

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||||||||||| |||||||| ||||| | | | || ||||| |||||||  ||
Db    517 AGATAGATGAGCTGGCTGACGTCTACACGGATGTTAGAAGACGATGGAAGGCAGAGATTT 576

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    577 TCATCTTCATGGGTGATTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 636

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||  ||||||| |||||||| || |||||||||| ||||||||||||||||||||
Db    637 ACATCCGTTTGAGGACAGACCCCAACTTTGTTTGGCTGATTGGGGACCAAGAGGACACCA 696

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||||||||||||||| | ||| |||||||||||||| ||  || | |||||| | || |
Db    697 CGGTCAAGAAGAGCACCAGCTGTGCCTATGACAGGATTGTCGTTCGCGGACAAGAGATAG 756

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          ||||||||  | ||| |  || |    |||||||| || ||||||||||   |||||
Db    757 TCAACTCTGTGGTTCCCCGCTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGT 816

Qy    601 CTGAATCGAAGGCCCTG---GATGTCAGCGACCACTTTCCAGTTCA 643
          |||||  | ||||||||   |||||||| ||||||||||||||| |
Db    817 CTGAAGAGGAGGCCCTGGATGATGTCAGTGACCACTTTCCAGTTGA 862
```

```
RESULT 12
E11687
LOCUS       E11687                  1208 bp    RNA     linear   PAT 29-SEP-1997
DEFINITION  cDNA encoding novel deoxyribonuclease(DNase) gamma which cut
            specifically linker site of chromatin DNA.
ACCESSION   E11687
VERSION     E11687.1  GI:22025323
KEYWORDS    JP 1996187079-A/1.
SOURCE      Rattus rattus.
  ORGANISM  Rattus rattus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1208)
  AUTHORS   Tanuma,Y.
  TITLE     NEW DEOXYRIBONUCLEASE
  JOURNAL   Patent: JP 1996187079-A 1 23-JUL-1996;
            TANUMA YASUKAZU
COMMENT     OS   Rattus rattus (rat)
            PN   JP 1996187079-A/1
            PD   23-JUL-1996
            PF   06-SEP-1995 JP 1995255647
            PR   06-SEP-1994 JP 94P   239518
            PI   TANUMA YASUKAZU
            PC   C12N9/16,C07K14/47,C07K16/40,C12N1/21,C12N15/09,(C12N1/21, PC
             C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .1208
            FT                   /organism='Rattus rattus'
            FT                   /tissue_type='Thymus, Spleen' FT
                /cell_type='Cell nucleus'
            FT   5'UTR           1. .11
            FT   CDS             12. .944
            FT                   /product='Novel deoxynbonuclease(DNase) gamma
            FT                   which cut
            FT                   specifically linker site of chromatin DNA' FT
             precursor_RNA   12. .86
            FT   3'UTR           945. .1208.
FEATURES             Location/Qualifiers
     source          1. .1208
                     /organism="Rattus rattus"
                     /db_xref="taxon:10117"
BASE COUNT      319 a    304 c    283 g    302 t
ORIGIN

  Query Match             71.7%;  Score 481.2;  DB 6;  Length 1208;
  Best Local Similarity   85.1%;  Pred. No. 3.5e-120;
  Matches  550;  Conservative    0;  Mismatches   93;  Indels    3;  Gaps    1;

Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||  ||||| ||||| ||||||||||||| || ||||| ||| ||| | |||  |
Db    217 ACAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAACTCACGAAGAAGCACGA 276

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          |||||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    277 CATACAACTACGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAGTATGCCT 336

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          | ||||| || || ||||| |||||||| ||    ||  |||||||||| ||||||||||
Db    337 TCCTCTACAAGGAGAAGCTGGTGTCTGTGAAGGCAAAATACCTCTACCATGACTATCAGG 396

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||  |||| ||||||||||||| |||||||||| |||||||| | |||| ||
Db    397 ATGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCA 456

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
```

```
Db    457 CTGCTGCCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCTGAAACCTCCGTTAAAG 516

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||||||||||| |||||||| ||||| | | | || ||||| ||||||| ||
Db    517 AGATAGATGAGCTGGCTGACGTCTACACGGATGTTAGAAGACGATGGAAGGCAGAGATTT 576

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    577 TCATCTTCATGGGTGATTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 636

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||  ||||||| |||||||| || ||||||||||| ||||||||||||||||||
Db    637 ACATCCGTTTGAGGACAGACCCCAACTTTGTTTGGCTGATTGGGGACCAAGAGGACACCA 696

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||||||||||||||  | ||| |||||||||||||  ||  || | |||||| | || |
Db    697 CGGTCAAGAAGAGCACCAGCTGTGCCTATGACAGGATTGTCGTTCGCGGACAAGAGATAG 756

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||||||||  | ||| |  || |    |||||||||| || ||||||||||||   |||||
Db    757 TCAACTCTGTGGTTCCCCGCTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGT 816

Qy    601 CTGAATCGAAGGCCCTG---GATGTCAGCGACCACTTTCCAGTTCA 643
          |||||  | ||||||||   ||||||||  ||||||||||||||||
Db    817 CTGAAGAGGAGGCCCTGGATGATGTCAGTGACCACTTTCCAGTTGA 862


RESULT 13
MMU76110
LOCUS       MMU76110                 941 bp    mRNA    linear   ROD 26-JAN-1999
DEFINITION  Mus musculus DNase gamma mRNA, complete cds.
ACCESSION   U76110
VERSION     U76110.1  GI:4098207
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 941)
  AUTHORS   Shiokawa,D., Hatanaka,T. and Tanuma,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-1996) Biochemistry, Science Univ. of Tokyo,
            Shinjuku-ku ichigaya funagawaracho, Tokyo 162, Japan
FEATURES             Location/Qualifiers
     source          1. .941
                     /organism="Mus musculus"
                     /strain="C57 black"
                     /db_xref="taxon:10090"
     CDS             7. .939
                     /function="apoptosis"
                     /note="deoxyribonuclease"
                     /codon_start=1
                     /product="DNase gamma"
                     /protein_id="AAD09222.1"
                     /db_xref="GI:4098208"
                     /translation="MSLHPASPRLASLLLFILALHDTLALRLCSFNVRSFGASKKENH
                     EAMDIIVKIIKRCDLILLMEIKDSSNNICPMLMEKLNGNSRRSTTYNYVISSRLGRNT
                     YKEQYAFVYKEKLVSVKTKYHYHDYQDGDTDVFSREPFVVWFHSPFTAVKDFVIVPLH
                     TTPETSVKEIDELVDVYTDVRSQWKTENFIFMGDFNAGCSYVPKKAWQNIRLRTDPKF
                     VWLIGDQEDTTVKKSTSCAYDRIVLCGQEIVNSVVPRSSGVFDFQKAYDLSEEEALDV
                     SDHFPVEFKLQSSRAFTNNRKSVSLKKRKKGNRS"
BASE COUNT      255 a    229 c    228 g    229 t
ORIGIN

  Query Match           70.8%;  Score 474.8;  DB 10;  Length 941;
  Best Local Similarity  82.9%;  Pred. No. 1.9e-118;
  Matches  542;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy      2 CAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAAC 61
          |||||  ||||| ||||| |||||||||||||| || |||||||| ||| | ||| |||
Db    213 CAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAATTCACGAAGAAGCACAAC 272

Qy     62 ATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCTT 121
          |||||||||||||||| ||||| |||||||||||||| || ||||| ||||||||||
Db    273 ATACAACTATGTGATTAGTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTATGCCTT 332

Qy    122 TCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGC 181
            |||| || || ||||| |||||||| ||    || |||| |||||| ||||||||||
Db    333 CGTCTACAAGGAGAAGCTGGTGTCTGTGAAGACAAAATACCACTACCATGACTATCAGGA 392

Qy    182 TGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACAC 241
          ||||||| |||| |||||||||||||| |||||||||| |||||||| || ||||  ||
Db    393 TGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTAC 452

Qy    242 CGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGA 301
          ||||||||||||||||||||||||||| ||||||| || || ||||| ||||||| |||
Db    453 TGCTGTCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTTAAAGA 512

Qy    302 GATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTT 361
          ||| |||||||||  |||||||||| ||||||| | | |  |||||  ||||||||||
Db    513 GATAGATGAGCTGGTCGATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTT 572

Qy    362 CATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGA 421
          ||| |||||||||| ||||| || ||||| ||||| |||||||||||||||||| || |
Db    573 CATCTTCATGGGTGATTTCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAA 632

Qy    422 CATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCAC 481
          ||| ||  ||||||||||||||||||| |||||||||| ||||||||||||||||||  ||
Db    633 CATTCGTTTGAGGACGGACCCCAAGTTTGTTTGGCTGATTGGGGACCAAGAGGACACTAC 692

Qy    482 GGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGT 541
          |||||||||||| || | ||| |||||||||||||| |||||| | |||||| | || ||
Db    693 GGTCAAGAAGAGTACCAGCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAGATAGT 752

Qy    542 CAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTC 601
          ||||||  |  | |||  |  || |    ||||||||| || ||||||||||    |||||
Db    753 CAACTCCGTGGTTCCCCGTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTC 812

Qy    602 TGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||   | |||||||||||||||| || ||||||||||| |   | | |  ||
Db    813 TGAGGAGGAGGCCCTGGATGTCAGTGATCACTTTCCAGTTGAGTTTAAGCTACA 866


RESULT 14
AF047355
LOCUS       AF047355                 1124 bp    mRNA    linear   ROD 16-SEP-1998
DEFINITION  Mus musculus liver and spleen DNase precursor (LSD) mRNA, complete
            cds.
ACCESSION   AF047355
VERSION     AF047355.1  GI:2905787
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1124)
  AUTHORS   Baron,W.F., Pan,C.Q., Spencer,S.A., Ryan,A.M., Lazarus,R.A. and
            Baker,K.P.
  TITLE     Cloning and characterization of an actin-resistant DNase I-like
            endonuclease secreted by macrophages
  JOURNAL   Gene 215 (2), 291-301 (1998)
  MEDLINE   98382522
   PUBMED   9714828
REFERENCE   2  (bases 1 to 1124)
  AUTHORS   Baron,W.F., Pan,C.Q., Spencer,S.A., Ryan,A.M., Lazarus,R.A. and
            Baker,K.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-1998) Molecular Biology, Genentech Inc., 1 DNA
            Way, South San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1. .1124
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     gene            1. .1124
                     /gene="LSD"
```

```
     CDS             173. .1105
                     /gene="LSD"
                     /note="actin resistant acivity; macrophage specific DNase
                     I; LS-DNase"
                     /codon_start=1
                     /product="liver and spleen DNase precursor"
                     /protein_id="AAC35753.1"
                     /db_xref="GI:2905788"
                     /translation="MSLHPASPRLASLLLFILALHDTLALRLCSFNVRSFGASKKENH
                     EAMDIIVKIIKRCDLILLMEIKDSSNNICPMLMEKLNGNSRRSTTYNYVISSRLGRNT
                     YKEQYAFVYKEKLVSVKTKYHYHDYQDGDTDVFSREPFVVWFHSPFTAVKDFVIVPLH
                     TTPETSVKEIDELVDVYTDVRSQWKTENFIFMGDFNAGCSYVPKKAWQNIRLRTDPKF
                     VWLIGDQEDTTVKKSTSCAYDRIVLCGQEIVNSVVPRSSGVFDFQKAYDLSEEEALDV
                     SDHFPVEFKLQSSRAFTNNRKSVSLKKRKKGNRS"
     sig_peptide     173. .247
                     /gene="LSD"
                     /evidence=not_experimental
     mat_peptide     248. .1102
                     /gene="LSD"
                     /product="liver and spleen DNase"
BASE COUNT      299 a    273 c    273 g    279 t
ORIGIN

  Query Match             70.8%;  Score 474.8;  DB 10;  Length 1124;
  Best Local Similarity   82.9%;  Pred. No. 1.9e-118;
  Matches  542;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy      2 CAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAAC 61
          |||||  |||| |||| ||||||||||||| || |||||||| ||| | ||| |||
Db    379 CAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAATTCACGAAGAAGCACAAC 438

Qy     62 ATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCTT 121
          |||||||||||||||||| ||||| |||||||||||||| || ||||| ||||||||||
Db    439 ATACAACTATGTGATTAGTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTATGCCTT 498

Qy    122 TCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGC 181
            |||| || || ||||| |||||||| ||   ||  |||| |||||| ||||||||||
Db    499 CGTCTACAAGGAGAAGCTGGTGTCTGTGAAGACAAAATACCACTACCATGACTATCAGGA 558

Qy    182 TGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACAC 241
          ||||||| |||| |||||||||||||| |||||||||| |||||||| || ||||  ||
Db    559 TGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTAC 618

Qy    242 CGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGA 301
          |||||||||||||||||||||||||||| ||||||| || || ||||| ||||||| |||
Db    619 TGCTGTCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTTAAAGA 678

Qy    302 GATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTT 361
          ||| ||||||||||  ||||||||||| ||||||| | | |  |||||  ||||||||||
Db    679 GATAGATGAGCTGGTCGATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTT 738

Qy    362 CATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGA 421
          ||| |||||||||| ||||| || ||||| ||||| |||||||||||||||||| || |
Db    739 CATCTTCATGGGTGATTTCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAA 798

Qy    422 CATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCAC 481
          ||| ||  |||||||||||||||||| |||||||||| ||||||||||||||||||| ||
Db    799 CATTCGTTTGAGGACGGACCCCAAGTTTGTTTGGCTGATTGGGGACCAAGAGGACACTAC 858

Qy    482 GGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGT 541
          |||||||||||| || | ||| |||||||||||||| |||||| | |||||| | || ||
Db    859 GGTCAAGAAGAGTACCAGCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAGATAGT 918

Qy    542 CAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTC 601
          |||||| |  | ||| |  || |     ||||||||| || ||||||||||     |||||
Db    919 CAACTCCGTGGTTCCCCGTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTC 978

Qy    602 TGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||   | ||||||||||||||| || |||||||||||  |   | | |  ||
Db    979 TGAGGAGGAGGCCCTGGATGTCAGTGATCACTTTCCAGTTGAGTTTAAGCTACA 1032


RESULT 15
BC012671
LOCUS       BC012671               2110 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, Similar to deoxyribonuclease 1-like 3, clone
            MGC:13854 IMAGE:4160709, mRNA, complete cds.
ACCESSION   BC012671
VERSION     BC012671.1  GI:15215118
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2110)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 18 Row: i Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 6681204.
FEATURES             Location/Qualifiers
     source          1. .2110
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="MGC:13854 IMAGE:4160709"
                     /tissue_type="Liver, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             151. .1083
                     /codon_start=1
                     /product="Similar to deoxyribonuclease 1-like 3"
                     /protein_id="AAH12671.1"
                     /db_xref="GI:15215119"
                     /translation="MSLHPASPRLASLLLFILALHDTLALRLCSFNVRSFGASKKENH
                     EAMDIIVKIIKRCDLILLMEIKDSSNNICPMLMEKLNGNSRRSTIYNYVISSRLGRNT
                     YKEQYAFVYKEKLVSVKTKYHYHDYQDGDTDVFSREPFVVWFHSPFTAVKDFVIVPLH
                     TTPETSVKEIDELVDVYTDVRSQWKTENFIFMGDFNAGCSYVPKKAWQNIRLRTDPKF
                     VWLIGDQEDTTVKKSTSCAYDRIVLCGQEIVNSVVPRSSGVFDFQKAYDLSEEEALDV
                     SDHFPVEFKLQSSRAFTNNRKSVSLKKRKKGNRS"
BASE COUNT      526 a    581 c    464 g    539 t
ORIGIN

  Query Match             70.5%;  Score 473.2;  DB 10;  Length 2110;
  Best Local Similarity   82.7%;  Pred. No. 5.5e-118;
  Matches  541;  Conservative    0;  Mismatches  113;  Indels    0;  Gaps    0;

Qy      2 CAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAAC 61
          |||||  |||| |||| ||||||||||||| || |||||||| ||| | ||| ||
Db    357 CAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAATTCACGAAGAAGCACAAT 416

Qy     62 ATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCTT 121
          |||||||||||||||||| ||||| |||||||||||||| || ||||| ||||||||||
Db    417 ATACAACTATGTGATTAGTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTATGCCTT 476

Qy    122 TCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGC 181
            |||| || || ||||| |||||||| ||   ||  |||| |||||| ||||||||||
```

```
Db    477 CGTCTACAAGGAGAAGCTGGTGTCTGTGAAGACAAAATACCACTACCATGACTATCAGGA 536

Qy    182 TGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACAC 241
          ||||||| |||| |||||||||||||| |||||||||| |||||||| || ||||  ||
Db    537 TGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTAC 596

Qy    242 CGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGA 301
           |||||||||||||||||||||||||| ||||||| || || ||||| ||||||| |||
Db    597 TGCTGTCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTTAAAGA 656

Qy    302 GATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTT 361
          ||| |||||||||| |||||||||| ||||||| | | |  |||||  ||||||||||
Db    657 GATAGATGAGCTGGTCGATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTT 716

Qy    362 CATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGA 421
          ||| |||||||||| ||||| || ||||| ||||| ||||||||||||||||| || |
Db    717 CATCTTCATGGGTGATTTCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAA 776

Qy    422 CATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCAC 481
          ||| ||  |||||||||||||||||| |||||||||| |||||||||||||||| ||
Db    777 CATTCGTTTGAGGACGGACCCCAAGTTTGTTTGGCTGATTGGGGACCAAGAGGACACTAC 836

Qy    482 GGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGT 541
          |||||||||||| || | ||| |||||||||||||| |||||| | |||||| | || ||
Db    837 GGTCAAGAAGAGTACCAGCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAGATAGT 896

Qy    542 CAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTC 601
          |||||| |  | ||| |  || |    |||||||| || ||||||||||    |||||
Db    897 CAACTCCGTGGTTCCCCGTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTC 956

Qy    602 TGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||   | |||||||||||||| || |||||||||||| |   | | |  ||
Db    957 TGAGGAGGAGGCCCTGGATGTCAGTGATCACTTTCCAGTTGAGTTTAAGCTACA 1010
```

Search completed: December 9, 2002, 11:30:57
Job time : 1421.41 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 08:59:14 ; Search time 193.384 Seconds
(without alignments)
7813.939 Million cell updates/sec

Title: US-09-905-114-3
Perfect score: 671
Sequence: 1 acaacaggatctgccccata.....atcatcatgaagaaccatga 671

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	671	100.0	671	24	AAD29089	Bovine 22kDa recom
	2	592	88.2	592	24	AAD29088	Bovine fertility a
	3	539	80.3	578	24	AAS17900	DNA encoding bovin
	4	527	78.5	1023	24	ABN96895	Gene #3393 used to
	5	527	78.5	1023	24	ABL61914	Colon adenocarcino
	6	527	78.5	1023	24	ABL64458	Stomach cancer rel
	7	527	78.5	1023	24	ABL64961	Lung cancer relate
	8	527	78.5	1079	18	AAT74076	Human LS-DNase nuc
	9	525.4	78.3	1108	24	ABK84429	Human cDNA differe
	10	518.4	77.3	556	24	AAS17901	Partail cDNA encod
	11	479.6	71.5	1208	17	AAT17733	Deoxyribonuclease,
	12	474.8	70.8	1124	18	AAT74083	Murine LS-DNase nu
	13	348.4	51.9	350	24	AAS17902	partial cDNA encod
	14	314.8	46.9	350	24	AAS17903	partial cDNA encod
	15	300.6	44.8	349	24	AAS17904	Partial cDNA encod
c	16	195.6	29.2	392	24	ABN94132	Gene #630 used to
c	17	195.6	29.2	392	24	ABL62246	Colon adenocarcino
	18	163.2	24.3	1039	15	AAQ54429	Human DNase I. Ho
	19	163.2	24.3	1039	22	ABA02661	Human DNase I 1.
	20	161.6	24.1	783	22	ABA02662	Human DNase I 2.
	21	161.6	24.1	858	22	ABA02663	Human DNase I 3.
	22	161.6	24.1	1000	21	AAA28433	DNA encoding STII
	23	161.6	24.1	1039	11	AAQ05512	Sequence encoding
	24	161.6	24.1	1039	15	AAQ62645	Human DNase gene.
	25	161.6	24.1	1548	22	ABA02702	Humanised HMFG-1 F
	26	161.6	24.1	1548	22	ABA02703	Humanised HMFG-1 F
	27	161.6	24.1	1548	22	ABA02705	Humanised HMFG-1 F
	28	161.6	24.1	1554	22	ABA02682	Humanised HMFG-1 F
	29	161.6	24.1	1554	22	ABA02683	Humanised HMFG-1 F
	30	161.6	24.1	1554	22	ABA02685	Humanised HMFG-1 F
	31	161.6	24.1	1557	22	ABA02704	Humanised HMFG-1 F
	32	161.6	24.1	1560	22	ABA02710	Humanised HMFG-1 F
	33	161.6	24.1	1560	22	ABA02711	Humanised HMFG-1 F
	34	161.6	24.1	1560	22	ABA02713	Humanised HMFG-1 F
	35	161.6	24.1	1560	22	ABA02714	Humanised HMFG-1 F
	36	161.6	24.1	1560	22	ABA02715	Humanised HMFG-1 F
	37	161.6	24.1	1560	22	ABA02716	Humanised HMFG-1 F
	38	161.6	24.1	1563	22	ABA02684	Humanised HMFG-1 F
	39	161.6	24.1	1566	22	ABA02706	Humanised HMFG-1 F
	40	161.6	24.1	1566	22	ABA02707	Humanised HMFG-1 F
	41	161.6	24.1	1566	22	ABA02709	Humanised HMFG-1 F
	42	161.6	24.1	1569	22	ABA02712	Humanised HMFG-1 F
	43	161.6	24.1	1575	22	ABA02708	Humanised HMFG-1 F
	44	161.6	24.1	1578	22	ABA02717	Humanised HMFG-1 F
	45	161.6	24.1	1578	22	ABA02718	Humanised HMFG-1 F

ALIGNMENTS

RESULT 1
AAD29089
ID AAD29089 standard; DNA; 671 BP.
XX
AC AAD29089;
XX
DT 16-MAY-2002 (first entry)
XX
DE Bovine 22kDa recombinant FAA (rFAA) coding sequence.
XX
KW Bovine; fertility associated antigen; FAA; sperm fertility; acrosome;
KW artificial insemination; plasma membrane; sperm cell; anti-infertility;
KW reproductive tract; ds.
XX
OS Bos sp.
XX
PN WO200206444-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US22097.
XX
PR 14-JUL-2000; 2000US-218140P.
XX
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
PI Zhang H, Ax RL, Bellin ME;
XX
DR WPI; 2002-171803/22.
XX
PT Novel amino acid and polynucleotide sequence for fertility associated
PT antigen useful for increasing stability of plasma membrane, acrosome

```
PT   and other portions of sperm cell, and increasing fertility of mammals
PT   -
XX
PS   Claim 1; Fig 5; 37pp; English.
XX
CC   The invention relates to bovine fertility associated antigen (FAA) and
CC   its nucleic acid sequence. FAA is useful for increasing the stability of
CC   the plasma membrane plus acrosome of a sperm cell and/or other portions
CC   of a sperm cell, in particular mammalian sperm cell such as a sperm cell
CC   from buffalo, cow, horse, mice, pig, sheep, human, avian sperm cell
CC   including turkey or chicken sperm cell. FAA is also useful for increasing
CC   the fertility of sperm from a mammal including buffalo, cow, horse, mice,
CC   pig, sheep or human, or avian (chicken or turkey). FAA is added to the
CC   suspension of sperm cells prior to preservation or prior to
CC   administration of the sperm cells to a mammal in artificial insemination.
CC   Alternatively, fertility of sperm from a mammal is increased by placing
CC   FAA into a female's reproductive tract prior to deposition of the sperm
CC   cells into the female's reproductive tract by copulation or artificial
CC   insemination. The fertility of the male mammal is increased by injecting
CC   FAA into the male mammal's reproductive tract. The present sequence is
CC   bovine 22kDa recombinant FAA (rFAA) coding sequence.
XX
SQ   Sequence 671 BP; 195 A; 166 C; 154 G; 156 T; 0 other;

  Query Match             100.0%;  Score 671;  DB 24;  Length 671;
  Best Local Similarity   100.0%;  Pred. No. 4e-202;
  Matches  671;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600

Qy     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCATCATG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCATCATG 660

Qy     661 AAGAACCATGA 671
```

```
          ||||||||||||
Db     661 AAGAACCATGA 671


RESULT 2
AAD29088
ID   AAD29088 standard; cDNA; 592 BP.
XX
AC   AAD29088;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Bovine fertility associated antigen (FAA) partial cDNA.
XX
KW   Bovine; fertility associated antigen; FAA; sperm fertility; acrosome;
KW   artificial insemination; plasma membrane; sperm cell; anti-infertility;
KW   reproductive tract; ss.
XX
OS   Bos sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..591
FT                   /*tag=  a
FT                   /product= "Bovine fertility associated antigen (FAA)"
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200206444-A2.
XX
PD   24-JAN-2002.
XX
PF   16-JUL-2001; 2001WO-US22097.
XX
PR   14-JUL-2000; 2000US-218140P.
XX
PA   (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
PI   Zhang H,  Ax RL,  Bellin ME;
XX
DR   WPI; 2002-171803/22.
DR   P-PSDB; AAE18281.
XX
PT   Novel amino acid and polynucleotide sequence for fertility associated
PT   antigen useful for increasing stability of plasma membrane, acrosome
PT   and other portions of sperm cell, and increasing fertility of mammals
PT   -
XX
PS   Claim 1; Fig 6; 37pp; English.
XX
CC   The invention relates to bovine fertility associated antigen (FAA) and
CC   its nucleic acid sequence. FAA is useful for increasing the stability of
CC   the plasma membrane plus acrosome of a sperm cell and/or other portions
CC   of a sperm cell, in particular mammalian sperm cell such as a sperm cell
CC   from buffalo, cow, horse, mice, pig, sheep, human, avian sperm cell
CC   including turkey or chicken sperm cell. FAA is also useful for increasing
CC   the fertility of sperm from a mammal including buffalo, cow, horse, mice,
CC   pig, sheep or human, or avian (chicken or turkey). FAA is added to the
CC   suspension of sperm cells prior to preservation or prior to
CC   administration of the sperm cells to a mammal in artificial insemination.
CC   Alternatively, fertility of sperm from a mammal is increased by placing
CC   FAA into a female's reproductive tract prior to deposition of the sperm
CC   cells into the female's reproductive tract by copulation or artificial
CC   insemination. The fertility of the male mammal is increased by injecting
CC   FAA into the male mammal's reproductive tract. The present sequence is
CC   bovine fertility associated antigen (FAA) partial cDNA.
XX
SQ   Sequence 592 BP; 170 A; 143 C; 142 G; 137 T; 0 other;

  Query Match             88.2%;  Score 592;  DB 24;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-177;
  Matches  592;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      27 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 86
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60

Qy     87 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120

Qy    147 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180

Qy    207 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240

Qy    267 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300

Qy    327 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360

Qy    387 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420

Qy    447 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480

Qy    507 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540

Qy    567 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592


RESULT 3
AAS17900
ID   AAS17900 standard; cDNA; 578 BP.
XX
AC   AAS17900;
XX
DT   26-MAR-2002  (first entry)
XX
DE   DNA encoding bovine fertility-associated antigen (FAA).
XX
KW   Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;
KW   reproductive fitness; single nucleotide polymorphism; SNP; bovine; cow;
KW   ss.
XX
OS   Bos taurus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..578
FT                   /*tag=  a
FT                   /product= "FAA"
FT                   /note= "Fertility-associated antigen"
FT                   /partial
FT                   /note= "No start or stop codon given"
FT                   /transl_except= (pos:289..300, aa:A)
FT                   /note= "This codon has an apparent 9 nucleotide
FT                          insertion"
FT                   /transl_except= (pos:325..326, aa:R)
FT                   /note= "This codon has an apparent 1 nucleotide
FT                           deletion which alters the reading frame"
FT                   /transl_except= (pos:576..578, aa:RLSESKAL)
XX
PN   WO200176529-A2.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-US10802.
XX
PR   07-APR-2000; 2000US-195225P.
XX
PA   (ZHAN/) ZHANG H M.
PA   (AXRL/) AX R L.
XX
PI   Zhang HM,  Ax RL;
XX
DR   WPI; 2002-066307/09.
DR   P-PSDB; AAU11516.
XX
PT   Novel human or bovine fertility-associated antigen useful for
PT   stabilising sperm cell acrosome and increasing fertility of a male  -
XX
PS   Claim 25; Fig 2A-E; 54pp; English.
XX
CC   The invention describes a novel isolated human fertility-associated
CC   antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC   for increasing the stability of a sperm cell acrosome and for increasing
CC   the fertility of a human male by administering FAA. Detecting single
CC   nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC   for assaying fertility by detecting the presence or absence of a FAA SNP
CC   in the mammal and correlating the presence or absence with the fertility
CC   of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC   specification. FAA SNP is also useful for determining the reproductive
CC   fitness of a human or bovine, by determining the nucleotide sequence of
CC   the FAA SNP, quantifying the fertility of more than one mammal containing
CC   the FAA SNP and correlating the frequency of the FAA SNP to the
CC   reproductive fitness of the mammals. FAA improves the integrity of sperm
CC   membranes and increases the capacitation of sperm derived from either
CC   fertile or infertile humans. Early genetic identification of infertility
CC   improves the prognosis for subsequent attempts at fertilisation and
CC   facilitates early intervention to determine whether the individual may
CC   benefit from fertility treatments to avoid costly and/or emotional
CC   problems with attempted inseminations. This sequence encodes the bovine
CC   fertility-associated antigen (FAA) described in the method of the
CC   invention.
XX
SQ   Sequence 578 BP; 167 A; 140 C; 138 G; 133 T; 0 other;

  Query Match             80.3%;  Score 539;  DB 24;  Length 578;
  Best Local Similarity   98.3%;  Pred. No. 2.6e-160;
  Matches  569;  Conservative    0;  Mismatches    0;  Indels   10;  Gaps    2;

Qy     27 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 86
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60

Qy     87 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120

Qy    147 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180

Qy    207 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240

Qy    267 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGA---------TGAGCTGGCT 317
          ||||||||||||||||||||||||||||||||||||||||||         ||||||||||
Db    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGAGCTGGCT 300

Qy    318 GATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGAC 377
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    301 GATGTCTACACAGATGTGAAACGTC-CTGGAATGCAGAGAATTTCATTTTCATGGGTGAC 359

Qy    378 TTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACG 437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    360 TTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACG 419

Qy    438 GACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACA 497
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 GACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACA 479

Qy    498 AACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCT 557
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 AACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCT 539

Qy    558 CAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
          |||||||||||||||||||||||||||||||||||||||
Db    540 CAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 578
```

RESULT 4
ABN96895
ID ABN96895 standard; DNA; 1023 BP.
XX
AC ABN96895;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3393 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 3393; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

Query Match 78.5%; Score 527; DB 24; Length 1023;
Best Local Similarity 87.8%; Pred. No. 2.2e-156;
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||||||||||||||||||||||||||||||||| ||| |||||||||| | |||||||||
Db    215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||||| |||||||||||||||||||||||||||| |||||||
Db    275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db    335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||| |||||||||||||||||||| || ||| |||
Db    395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||| |||||||||||| |
Db    455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db    515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||
Db    635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||| ||||| || |||||||||| ||||||||||||||| | || |
Db    695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db    755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||||| |   | | |  ||
Db    815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869
```

RESULT 5
ABL61914
ID ABL61914 standard; DNA; 1023 BP.
XX
AC ABL61914;
XX
DT 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:251.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 251; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

Query Match 78.5%; Score 527; DB 24; Length 1023;
Best Local Similarity 87.8%; Pred. No. 2.2e-156;
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||||||| ||||||||||||||||||||||||| |||||||
Db    275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| ||||||| ||     || || | |||||| ||||||||||
Db    335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db    395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||| ||||||||||||  |
Db    455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db    515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||||
Db    635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db    695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db    755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||| |   | | |  ||
Db    815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869
```

RESULT 6
ABL64458
ID ABL64458 standard; DNA; 1023 BP.
XX
AC ABL64458;
XX
DT 15-MAY-2002 (first entry)
XX
DE Stomach cancer related gene sequence SEQ ID NO:2795.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 2795; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

Query Match 78.5%; Score 527; DB 24; Length 1023;
Best Local Similarity 87.8%; Pred. No. 2.2e-156;
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||| |||||||||||||||||||||||||| |||||||
Db    275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db    335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db    395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||| ||||||||||||  |
Db    455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || ||||||| |||||||| || |||||||
Db    515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          ||||||| ||||||| ||||||| ||| |||||||||||||||||||||||||||||||
Db    635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db    695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| |||||||||||||||| | ||
Db    755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||| |   | | |  ||
Db    815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869
```

RESULT 7
ABL64961
ID ABL64961 standard; DNA; 1023 BP.
XX
AC ABL64961;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3298.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX

PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 3298; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.

XX
SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

Query Match 78.5%; Score 527; DB 24; Length 1023;
Best Local Similarity 87.8%; Pred. No. 2.2e-156;
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          |||||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||||| |||||||||||||||||||||||||||| |||||||
Db    275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db    335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||| |||||||||||||||||||| || ||| |||
Db    395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||| ||||||||||||  |
Db    455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || ||||||| |||||||| || |||||||
Db    515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||
Db    635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||||| ||||| || |||||||||||| |||||||||||||| | || |
Db    695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| |||||||||||||||| | ||
Db    755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||||| |   | | |  ||
Db    815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869
```

RESULT 8
AAT74076
ID AAT74076 standard; cDNA; 1079 BP.
XX
AC AAT74076;
XX
DT 29-JAN-1998 (first entry)
XX
DE Human LS-DNase nucleotide sequence.
XX
KW DNase; actin; DNA viscoelasticity; systemic lupus erythematosus;
KW cystic fibrosis; meningitis; pulmonary disease; gene therapy;
KW recombinant preparation; DNA-hydrolytic activity; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 71..988
FT /*tag= a
FT sig_peptide 71..130
FT /*tag= b

```
FT   mat_peptide      131..985
FT                    /*tag= c
FT                    /product= LS-DNase
XX
PN   WO9728266-A1.
XX
PD   07-AUG-1997.
XX
PF   03-FEB-1997;   97WO-US01506.
XX
PR   05-FEB-1996;   96US-0597078.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Baron WF;
XX
DR   WPI; 1997-402619/37.
DR   P-PSDB; AAW23542.
XX
PT   New isolated LS-DNase which is resistant to inhibition by actin -
PT   used for the treatment of e.g systemic lupus erythematosus and
PT   pulmonary diseases and disorders such as cystic fibrosis
XX
PS   Claim 1; Fig 1; 32pp; English.
XX
CC   This DNA sequence encodes a novel human LS-DNase (AAW23542), which has
CC   DNA-hydrolytic activity but is resistant to inhibition by actin. It can
CC   be used for reducing the viscoelasticity or viscous consistency of
CC   DNA-containing material, as well as treatment of systemic lupus
CC   erythematosus or pulmonary disease of disorders such as cystic fibrosis,
CC   bronchitis, pneumonia, bronchiectasis, emphysema, asthma, tuberculosis
CC   or fungal infections. It can also be used for adjunctive treatment of
CC   abscesses or severe closed-space infections in conditions such as
CC   empyema, meningitis, peritonitis, sinusitis, otitis, periodontitis,
CC   pericarditis, pancreatitis, cholelithiasis, endocarditis and
CC   septic arthritis, as well as in topical treatments in a variety of
CC   inflammatory and infected lesions, e.g. of the skin and mucosal
CC   membranes, surgical wounds, ulcerative lesions and burns. LS-DNase may
CC   improve the efficacy of antibiotics used in the treatment of such
CC   infections (e.g. gentamicin activity is markedly reduced by reversible
CC   binding to intact DNA). It can also be used for the treatment of other
CC   non-infected conditions in which there is an accumulation of cellular
CC   debris that includes cellular DNA, such as pyelonephritis and
CC   tubulo-interstitial kidney disease. It can also be used in in vitro
CC   diagnostic assays of a viscous material (e.g. sputum) from a patient.
CC   The anti-LS-DNase antibodies can be used for detection and purification
CC   of LS-DNase.
XX
SQ   Sequence 1079 BP; 330 A; 260 C; 252 G; 237 T; 0 other;

  Query Match             78.5%;  Score 527;  DB 18;  Length 1079;
  Best Local Similarity   87.8%;  Pred. No. 2.3e-156;
  Matches  575;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
           |||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db     261 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 320

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
           | |||||||||||||||||||| |||||||||||||||||||||||||||| |||||||
Db     321 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 380

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
           ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db     381 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 440

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
           |||||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db     441 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 500

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
           | |||||||| |||||||||||| |||||||||||||||||| |||||||||||| |
Db     501 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 560

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
           |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db     561 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 620

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     621 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 680

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
           |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||
Db     681 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 740

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
           |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db     741 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 800

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
           |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db     801 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 860

Qy     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
           |||||  | |||||||||||||||||||||||||||||| |   | | |  ||
Db     861 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 915


RESULT 9
ABK84429
ID   ABK84429 standard; cDNA; 1108 BP.
XX
AC   ABK84429;
XX
DT   14-AUG-2002  (first entry)
XX
DE   Human cDNA differentially expressed in granulocytic cells #1000.
XX
KW   Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW   viral infection; parasitic infection; protozoal infection;
KW   fungal infection; sterile inflammatory disease; psoriasis;
KW   rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW   cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW   adult respiratory distress syndrome; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; periodontal disease; .
KW   granulocyte activation; chronic inflammation; allergy.
XX
OS   Homo sapiens.
XX
PN   WO200228999-A2.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US30821.
XX
PR   03-OCT-2000; 2000US-237189P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Beazer-Barclay Y,  Weissman SM,  Yamaga S,  Vockley J;
XX
DR   WPI; 2002-435328/46.
XX
PT   Detecting granulocyte activation by detecting differential expression
PT   of genes associated with granulocyte activation, which serves as
PT   diagnostic markers that is useful for monitoring disease states and
PT   drug toxicity  -
XX
PS   Claim 1; SEQ ID No 1000; 114pp; English.
XX
CC   The invention relates to detecting (M1) granulocyte (GC) activation
CC   (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC   DNA chip analysis as given in the specification, and comparing
CC   the expression level to an expression level in an unactivated
CC   GC, where differential expression of Gs is indicative of GCA.
```

```
CC   Also included are modulating (M2) GA by contacting GC with an agent
CC   that alters the expression of at least one gene in Gs; (2) screening (M3)
CC   for an agent capable of modulating GCA or an inflammation (especially
CC   chronic) in a tissue, an allergic response in a subject, exposure of a
CC   subject to a pathogen or sterile inflammatory disease using the
CC   gene expression profile; (3) detecting (M4) an inflammation (especially
CC   chronic) in a tissue, an allergic response in a subject, exposure of a
CC   subject to a pathogen or sterile inflammatory disease, by detecting the
CC   level of expression in a sample of the tissue of gene(s) from Gs, where
CC   the level of expression of the gene is indicative of inflammation;
CC   (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC   an allergic response in a subject, exposure of a subject to a pathogen
CC   or sterile inflammatory disease, by contacting a tissue having
CC   inflammation with an agent that modulates the expression of gene(s)
CC   from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC   modulating GA; M3 is useful for screening an agent capable of modulating
CC   GCA preferably in an inflammation in a tissue; M4 is useful for
CC   detecting an inflammation (especially chronic) in a tissue, an allergic
CC   response in a subject, exposure of a subject to a pathogen or sterile
CC   inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC   glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC   reperfusion injury, ARDS, adult respiratory distress syndrome,
CC   inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC   periodontal disease; also bacterial infection, viral infection,
CC   parasitic infection, protozoal infection, fungal infection and M5 is
CC   useful for treating one of the above conditions. The present
CC   sequence represents a gene differentially expressed in granulocytes.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1108 BP; 329 A; 261 C; 271 G; 247 T; 0 other;

  Query Match             78.3%;  Score 525.4;  DB 24;  Length 1108;
  Best Local Similarity   87.6%;  Pred. No. 7.5e-156;
  Matches  574;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
Db     283 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 342

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
Db     343 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAAAACATATAAAGAACAATATGCCT 402

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
Db     403 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 462

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
Db     463 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 522

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
Db     523 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 582

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
Db     583 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 642

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
Db     643 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 702

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
Db     703 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 762

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
Db     763 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 822

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
Db     823 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 882

Qy     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
Db     883 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 937


RESULT 10
AAS17901
ID   AAS17901 standard; cDNA; 556 BP.
XX
AC   AAS17901;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Partail cDNA encoding human fertility-associated antigen (FAA), HC1.
XX
KW   Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;
KW   reproductive fitness; single nucleotide polymorphism; SNP; human;
KW   HC1; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..555
FT                   /*tag=  a
FT                   /product= "Human FAA"
FT                   /note= "Human fertility-associated antigen, HC1"
FT                   /partial
FT                   /note= "No start or stop codon given"
FT                   /transl_except= (pos:1..3, aa:KVIKRCDIILVMEIKDS)
FT                   /transl_except= (pos: 253..255, aa:S)
XX
PN   WO200176529-A2.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-US10802.
XX
PR   07-APR-2000; 2000US-195225P.
XX
PA   (ZHAN/) ZHANG H M.
PA   (AXRL/) AX R L.
XX
PI   Zhang HM,  Ax RL;
XX
DR   WPI; 2002-066307/09.
DR   P-PSDB; AAU11517.
XX
PT   Novel human or bovine fertility-associated antigen useful for
PT   stabilising sperm cell acrosome and increasing fertility of a male  -
XX
PS   Claim 1; Fig 2A-E; 54pp; English.
XX
CC   The invention describes a novel isolated human fertility-associated
CC   antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC   for increasing the stability of a sperm cell acrosome and for increasing
CC   the fertility of a human male by administering FAA. Detecting single
CC   nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC   for assaying fertility by detecting the presence or absence of a FAA SNP
CC   in the mammal and correlating the presence or absence with the fertility
CC   of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC   specification. FAA SNP is also useful for determining the reproductive
CC   fitness of a human or bovine, by determining the nucleotide sequence of
CC   the FAA SNP, quantifying the fertility of more than one mammal containing
CC   the FAA SNP and correlating the frequency of the FAA SNP to the
CC   reproductive fitness of the mammals. FAA improves the integrity of sperm
CC   membranes and increases the capacitation of sperm derived from either
CC   fertile or infertile humans. Early genetic identification of infertility
CC   improves the prognosis for subsequent attempts at fertilisation and
```

```
CC   facilitates early intervention to determine whether the individual may
CC   benefit from fertility treatments to avoid costly and/or emotional
CC   problems with attempted inseminations. This sequence encodes the human
CC   fertility-associated antigen (FAA), HC1, one of 4 partial prostate cDNA
CC   clones of the novel FAA gene described in the method of the invention.
XX
SQ   Sequence 556 BP; 159 A; 130 C; 142 G; 125 T; 0 other;

  Query Match             77.3%;  Score 518.4;  DB 24;  Length 556;
  Best Local Similarity   96.2%;  Pred. No. 8.5e-154;
  Matches  531;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
           |||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db       5 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 64

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
           | ||||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db      65 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 124

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
           ||||||| || |||||||| |||||||| ||    || || | |||||| |||||||||
Db     125 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 184

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185 ATGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 244

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     245 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 304

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     305 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 364

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     365 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 424

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     425 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 484

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     485 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 544

Qy     541 TCAACTCTGGTG 552
           |||||||| ||
Db     545 TCAACTCTGTTG 556


RESULT 11
AAT17733
ID   AAT17733 standard; cDNA; 1208 BP.
XX
AC   AAT17733;
XX
DT   17-OCT-1996  (first entry)
XX
DE   Deoxyribonuclease, DNase-gamma cDNA.
XX
KW   Deoxyribonuclease; DNase-gammma; selective cleavage; linker;
KW   chromatin DNA; cell nucleus; 3'-hydroxy; 5'-phosphate; apoptosis;
KW   carcinogenesis; AIDS development; diagnosis; prevention;
KW   treatment; cancer; autoimmune disease; viral infections; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             12..944
FT                   /*tag=  a
FT                   /note= "GAT codon comprising bases 837-839 has
FT                         no corresponding amino acid in AAR94021"
XX
PN   WO9607735-A1.
XX
PD   14-MAR-1996.
XX
PF   06-SEP-1995;   95WO-JP01775.
XX
PR   06-SEP-1994;   94JP-0239518.
XX
PA   (TANU/) TANUMA S.
XX
PI   Tanuma S;
XX
DR   WPI; 1996-171610/17.
DR   P-PSDB; AAR94021.
XX
PT   DNase which selectively cuts linker region of chromatin DNA  -
PT   useful in diagnosis, treatment and prevention of cancer, autoimmune
PT   diseases, viral infections, etc.
XX
PS   Claim 4; Pages 55-57; 75pp; Japanese.
XX
CC   The present sequence encodes the deoxyribonuclease, DNase-gammma,
CC   which is able to selectively cleave the linker part of chromatin
CC   DNA. DNase-gamma has a mol. wt. of 32000 by SDS-PAGE, or 30000 by
CC   gel filtration, an optimum pH of 5.6, is found in the cell nucleus,
CC   is not dependent on divalent cations, has a Zn(2+) inhibition
CC   IC(50) > 1mM and leaves 3'-OH and 5'-P termini after cleavage. The
CC   DNase can be used at a mol. level to resolve the control
CC   mechanisms of carcinogenesis, the autoimmune system and AIDS
CC   development. It can also be used in the diagnosis, prevention and
CC   treatment of cancer, autoimmune disease, viral infections, etc.,
CC   and as an apoptosis control agent.
XX
SQ   Sequence 1208 BP; 319 A; 304 C; 282 G; 303 T; 0 other;

  Query Match             71.5%;  Score 479.6;  DB 17;  Length 1208;
  Best Local Similarity   85.0%;  Pred. No. 2.6e-141;
  Matches  549;  Conservative    0;  Mismatches   94;  Indels    3;  Gaps    1;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
           ||||||  ||||| ||||| |||||||||||||| || ||||| ||| ||| | |||  |
Db     217 ACAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAACTCACGAAGAAGCACGA 276

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
           |||||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||
Db     277 CATACAACTACGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAGTATGCCT 336

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
           | ||||| || || ||||| |||||||| ||    ||  |||||||||||| ||||||||||
Db     337 TCCTCTACAAGGAGAAGCTGGTGTCTGTGAAGGCAAAATACCTCTACCATGACTATCAGG 396

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
           |||||||  |||| ||||||||||||| |||||||||| |||||||| | |||| ||
Db     397 ATGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCAGGCGCCCTTCA 456

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
           | |||| ||||||||||||||||||||| ||||||| || ||||| || || |||| ||
Db     457 CTGCTGCCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCTGAAACCTCTGTTAAAG 516

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
           |||| |||||||||||||| |||||||| ||||| | | | || ||||| ||||||| ||
Db     517 AGATAGATGAGCTGGCTGACGTCTACACGGATGTTAGAAGACGATGGAAGGCAGAGATTT 576

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
           |||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     577 TCATCTTCATGGGTGATTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 636

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
           |||||||  ||||||| |||||||| || |||||||||| ||||||||||||||||||||
Db     637 ACATCCGTTTGAGGACAGACCCCAACTTTGTTTGGCTGATTGGGGACCAAGAGGACACCA 696
```

```
Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          ||||||||||||||||| | ||| ||||||||||||||| ||  || | |||||| | || |
Db    697 CGGTCAAGAAGAGCACCAGCTGTGCCTATGACAGGATTGTCGTTCGCGGACAAGAGATAG 756

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||||||||  | ||| |  || |    ||||||||| || ||||||||||    |||||
Db    757 TCAACTCTGTGGTTCCCCGCTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGAGTTGT 816

Qy    601 CTGAATCGAAGGCCCTG---GATGTCAGCGACCACTTTCCAGTTCA 643
          |||||  | ||||||||   |||||||.|||||||||||||| |
Db    817 CTGAAGAGGAGGCCCTGGATGATGTCAGTGACCACTTTCCAGTTGA 862
```

RESULT 12
AAT74083
ID AAT74083 standard; cDNA; 1124 BP.
XX
AC AAT74083;
XX
DT 29-JAN-1998 (first entry)
XX
DE Murine LS-DNase nucleotide sequence.
XX
KW DNase; actin; DNA viscoelasticity; systemic lupus erythematosus;
KW cystic fibrosis; meningitis; pulmonary disease; gene therapy;
KW recombinant preparation; DNA-hydrolytic activity; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 173..1105
FT /*tag= a
FT sig_peptide 173..247
FT /*tag= b
FT /note= "putative"
FT mat_peptide 248..1105
FT /*tag= c
FT /product= LS-DNase
XX
PN WO9728266-A1.
XX
PD 07-AUG-1997.
XX
PF 03-FEB-1997; 97WO-US01506.
XX
PR 05-FEB-1996; 96US-0597078.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Baron WF;
XX
DR WPI; 1997-402619/37.
XX
PT New isolated LS-DNase which is resistant to inhibition by actin -
PT used for the treatment of e.g systemic lupus erythematosus and
PT pulmonary diseases and disorders such as cystic fibrosis
XX
PS Example 6; Fig 3; 32pp; English.
XX
CC This DNA sequence encodes a variant LS-DNase, which has DNA-hydrolytic
CC activity but is resistant to inhibition by actin. A 649 base pair
CC fragment of the coding sequence of cloned human LS-DNase (AAT74076) was
CC used to screen a murine liver cDNA library and obtain this nucleotide.
CC Murine mature LS-DNase is 84% identical to human LS-DNase. A portion of
CC the present sequence was used in Northern blot assays of various murine
CC tissues. Expression of murine LS-DNase mRNA was found to be highest in
CC the liver and spleen. Human LS-DNase can be used for reducing the
CC viscoelasticity or viscous consistency of DNA-containing material, as
CC well as treatment of systemic lupus erythematosus or pulmonary disease
CC disorders such as cystic fibrosis. It can also be used for adjunctive
CC treatment of abscesses or severe closed-space infections, e.g.
CC meningitis. It may improve the efficacy of antibiotics used in the
CC treatment of such infections (e.g. gentamicin activity is markedly
CC reduced by reversible binding to intact DNA). It can also be used in in
CC vitro diagnostic assays of a viscous material (e.g. sputum) from a
CC patient. The anti-LS-DNase antibodies can be used for detection and
CC purification of LS-DNase.
XX
SQ Sequence 1124 BP; 299 A; 273 C; 273 G; 279 T; 0 other;

Query Match 70.8%; Score 474.8; DB 18; Length 1124;
Best Local Similarity 82.9%; Pred. No. 8.2e-140;
Matches 542; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```
Qy      2 CAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAAC 61
          |||||  |||||  |||||  |||||||||||||| || |||||||| ||| | ||| |||
Db    379 CAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAATTCACGAAGAAGCACAAC 438

Qy     62 ATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCTT 121
          |||||||||||||||||| ||||| |||||||||||||| || ||||| ||||||||||
Db    439 ATACAACTATGTGATTAGTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTATGCCTT 498

Qy    122 TCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGC 181
            |||| || || ||||| |||||||| ||   ||  |||| |||||| ||||||||||
Db    499 CGTCTACAAGGAGAAGCTGGTGTCTGTGAAGACAAAATACCACTACCATGACTATCAGGA 558

Qy    182 TGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACAC 241
          ||||||| |||| |||||||||||||| |||||||||| |||||||| || ||||  ||
Db    559 TGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTAC 618

Qy    242 CGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGA 301
          |||||||||||||||||||||||||||||| ||||||| || || ||||| ||||||| |||
Db    619 TGCTGTCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTTAAAGA 678

Qy    302 GATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTT 361
          ||| ||||||||||  |||||||||| ||||||| | | |  |||||  ||||||||||
Db    679 GATAGATGAGCTGGTCGATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTT 738

Qy    362 CATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGA 421
          ||| |||||||||| ||||| || ||||| ||||| |||||||||||||||||||| || |
Db    739 CATCTTCATGGGTGATTTCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAA 798

Qy    422 CATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCAC 481
          ||| ||  |||||||||||||||||| |||||||||| |||||||||||||||| ||
Db    799 CATTCGTTTGAGGACGGACCCCAAGTTTGTTTGGCTGATTGGGGACCAAGAGGACACTAC 858

Qy    482 GGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGT 541
          |||||||||||| || | ||| |||||||||||||| |||||| | |||||| | || ||
Db    859 GGTCAAGAAGAGTACCAGCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAGATAGT 918

Qy    542 CAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTC 601
          |||||| |  | ||| |  || |    ||||||||| || ||||||||||    |||||
Db    919 CAACTCCGTGGTTCCCCGTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTC 978

Qy    602 TGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||   | |||||||||||||||| || |||||||||||| |   | | |  ||
Db    979 TGAGGAGGAGGCCCTGGATGTCAGTGATCACTTTCCAGTTGAGTTTAAGCTACA 1032
```

RESULT 13
AAS17902
ID AAS17902 standard; cDNA; 350 BP.
XX
AC AAS17902;
XX
DT 26-MAR-2002 (first entry)
XX
DE partial cDNA encoding human fertility-associated antigen (FAA), HC2.
XX
KW Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;
KW reproductive fitness; single nucleotide polymorphism; SNP; human;
KW HC2; ss.
XX
OS Homo sapiens.
XX

```
FH   Key             Location/Qualifiers
FT   CDS             1..349
FT                   /*tag=  a
FT                   /product= "Human FAA"
FT                   /note= "Human fertility-associated antigen, HC2"
FT                   /partial
FT                   /note= "No start or stop codon given"
XX
PN   WO200176529-A2.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-US10802.
XX
PR   07-APR-2000; 2000US-195225P.
XX
PA   (ZHAN/) ZHANG H M.
PA   (AXRL/) AX R L.
XX
PI   Zhang HM,  Ax RL;
XX
DR   WPI; 2002-066307/09.
DR   P-PSDB; AAU11518.
XX
PT   Novel human or bovine fertility-associated antigen useful for
PT   stabilising sperm cell acrosome and increasing fertility of a male  -
XX
PS   Claim 1; Fig 2A-E; 54pp; English.
XX
CC   The invention describes a novel isolated human fertility-associated
CC   antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC   for increasing the stability of a sperm cell acrosome and for increasing
CC   the fertility of a human male by administering FAA. Detecting single
CC   nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC   for assaying fertility by detecting the presence or absence of a FAA SNP
CC   in the mammal and correlating the presence or absence with the fertility
CC   of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC   specification. FAA SNP is also useful for determining the reproductive
CC   fitness of a human or bovine, by determining the nucleotide sequence of
CC   the FAA SNP, quantifying the fertility of more than one mammal containing
CC   the FAA SNP and correlating the frequency of the FAA SNP to the
CC   reproductive fitness of the mammals. FAA improves the integrity of sperm
CC   membranes and increases the capacitation of sperm derived from either
CC   fertile or infertile humans. Early genetic identification of infertility
CC   improves the prognosis for subsequent attempts at fertilisation and
CC   facilitates early intervention to determine whether the individual may
CC   benefit from fertility treatments to avoid costly and/or emotional
CC   problems with attempted inseminations. This sequence encodes the human
CC   fertility-associated antigen (FAA), HC2, one of 4 partial prostate cDNA
CC   clones of the novel FAA gene described in the method of the invention.
XX
SQ   Sequence 350 BP; 88 A; 91 C; 93 G; 78 T; 0 other;

  Query Match             51.9%;  Score 348.4;  DB 24;  Length 350;
  Best Local Similarity   99.7%;  Pred. No. 5.1e-100;
  Matches  349;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    203 CAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGAT 262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGAT 60

Qy    263 TGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGT 322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGT 120

Qy    323 CTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAA 382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAA 180

Qy    383 TGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCC 442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCC 240

Qy    443 CAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTG 502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTG 300

Qy    503 CGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTG 552
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    301 CGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGTTG 350


RESULT 14
AAS17903
ID   AAS17903 standard; cDNA; 350 BP.
XX
AC   AAS17903;
XX
DT   26-MAR-2002  (first entry)
XX
DE   partial cDNA encoding human fertility-associated antigen (FAA), HC3.
XX
KW   Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;
KW   reproductive fitness; single nucleotide polymorphism; SNP; human;
KW   HC3; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..349
FT                   /*tag=  a
FT                   /product= "Human FAA"
FT                   /note= "Human fertility associated antigen, HC3"
FT                   /partial
FT                   /note= "No start or stop codon given"
FT                   /transl_except= (pos:251..253, aa: Arg)
XX
PN   WO200176529-A2.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-US10802.
XX
PR   07-APR-2000; 2000US-195225P.
XX
PA   (ZHAN/) ZHANG H M.
PA   (AXRL/) AX R L.
XX
PI   Zhang HM,  Ax RL;
XX
DR   WPI; 2002-066307/09.
DR   P-PSDB; AAU11519.
XX
PT   Novel human or bovine fertility-associated antigen useful for
PT   stabilising sperm cell acrosome and increasing fertility of a male  -
XX
PS   Claim 1; Fig 2A-E; 54pp; English.
XX
CC   The invention describes a novel isolated human fertility-associated
CC   antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC   for increasing the stability of a sperm cell acrosome and for increasing
CC   the fertility of a human male by administering FAA. Detecting single
CC   nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC   for assaying fertility by detecting the presence or absence of a FAA SNP
CC   in the mammal and correlating the presence or absence with the fertility
CC   of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC   specification. FAA SNP is also useful for determining the reproductive
CC   fitness of a human or bovine, by determining the nucleotide sequence of
CC   the FAA SNP, quantifying the fertility of more than one mammal containing
CC   the FAA SNP and correlating the frequency of the FAA SNP to the
CC   reproductive fitness of the mammals. FAA improves the integrity of sperm
CC   membranes and increases the capacitation of sperm derived from either
CC   fertile or infertile humans. Early genetic identification of infertility
CC   improves the prognosis for subsequent attempts at fertilisation and
CC   facilitates early intervention to determine whether the individual may
CC   benefit from fertility treatments to avoid costly and/or emotional
```

```
CC   problems with attempted inseminations. This sequence encodes the human
CC   fertility-associated antigen (FAA), HC3, one of 4 partial prostate cDNA
CC   clones of the novel FAA gene described in the method of the invention.
XX
SQ   Sequence 350 BP; 89 A; 93 C; 94 G; 74 T; 0 other;

  Query Match             46.9%;  Score 314.8;  DB 24;  Length 350;
  Best Local Similarity   93.7%;  Pred. No. 2.3e-89;
  Matches  328;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy     203 CAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGAT 262
           |||||| |||||||||||||||||||| || ||| |||| |||||||| ||||||||||
Db       1 CAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGAT 60

Qy     263 TGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGT 322
           | |||||||||||||||||| |||||||||||||  ||||| |||||| ||| ||| ||
Db      61 TATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGT 120

Qy     323 CTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAA 382
           |||||| || |||||||  |||||||| || ||||||||||||||||||||||||||||
Db     121 CTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAA 180

Qy     383 TGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCC 442
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     181 TGCTGGCTGCAGCTACGTCCCCAAGAAGGCCCGGAAGGACATCCGCCTGAGGACGGACCC 240

Qy     443 CAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTG 502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTG 300

Qy     503 CGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTG 552
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     301 CGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGTTG 350


RESULT 15
AAS17904
ID   AAS17904 standard; cDNA; 349 BP.
XX
AC   AAS17904;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Partial cDNA encoding human fertility-associated antigen (FAA), HC4.
XX
KW   Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;
KW   reproductive fitness; single nucleotide polymorphism; SNP; human;
KW   HC4; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..348
FT                   /*tag= a
FT                   /product= "Human FAA"
FT                   /note= "Human fertility-associated antigen, HC4"
FT                   /partial
FT                   /note= "No start or stop codon given"
FT                   /transl_except= (pos: 11..12, aa:Phe)
FT                   /note= "This codon has an apparent 1 nucleotide
FT                          deletion which alters the reading frame"
XX
PN   WO200176529-A2.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-US10802.
XX
PR   07-APR-2000; 2000US-195225P.
XX
PA   (ZHAN/) ZHANG H M.
PA   (AXRL/) AX R L.
XX
PI   Zhang HM,  Ax RL;
XX
DR   WPI; 2002-066307/09.
DR   P-PSDB; AAU11520.
XX
PT   Novel human or bovine fertility-associated antigen useful for
PT   stabilising sperm cell acrosome and increasing fertility of a male  -
XX
PS   Claim 1; Fig 2A-E; 54pp; English.
XX
CC   The invention describes a novel isolated human fertility-associated
CC   antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC   for increasing the stability of a sperm cell acrosome and for increasing
CC   the fertility of a human male by administering FAA. Detecting single
CC   nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC   for assaying fertility by detecting the presence or absence of a FAA SNP
CC   in the mammal and correlating the presence or absence with the fertility
CC   of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC   specification. FAA SNP is also useful for determining the reproductive
CC   fitness of a human or bovine, by determining the nucleotide sequence of
CC   the FAA SNP, quantifying the fertility of more than one mammal containing
CC   the FAA SNP and correlating the frequency of the FAA SNP to the
CC   reproductive fitness of the mammals. FAA improves the integrity of sperm
CC   membranes and increases the capacitation of sperm derived from either
CC   fertile or infertile humans. Early genetic identification of infertility
CC   improves the prognosis for subsequent attempts at fertilisation and
CC   facilitates early intervention to determine whether the individual may
CC   benefit from fertility treatments to avoid costly and/or emotional
CC   problems with attempted inseminations. This sequence encodes the human
CC   fertility-associated antigen (FAA), HC4, one of 4 partial prostate cDNA
CC   clones of the novel FAA gene described in the method of the invention.
XX
SQ   Sequence 349 BP; 89 A; 91 C; 93 G; 76 T; 0 other;

  Query Match             44.8%;  Score 300.6;  DB 24;  Length 349;
  Best Local Similarity   91.6%;  Pred. No. 7.3e-85;
  Matches  318;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy     206 GGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGT 265
           || | || |||||||||||||||| || ||| |||| |||||||| |||||||||||| |
Db       3 GGGAGCCCTTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAAGACTTCGTGATTAT 62

Qy     266 CCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTA 325
           |||||||||||||||||| |||||||||||||  ||||| |||||| ||| ||| |||||
Db      63 CCCCCTGCACACCACCCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTA 122

Qy     326 CACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGC 385
           ||| || |||||||  |||||||| || ||||||||||||||||||||||||||||||||
Db     123 CACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGC 182

Qy     386 TGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAA 445
           |||||||||||||||||||||||||||||| |||||||| ||||||| |||||||
Db     183 CGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAG 242

Qy     446 GTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGC 505
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GTTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGC 302

Qy     506 CTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTG 552
           |||||||||||||||||||||||||||||||||||||||||||||| ||
Db     303 CTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGTTG 349


Search completed: December  9, 2002, 09:18:03
Job time : 201.384 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:20:03 ; Search time 1496.6 Seconds
(without alignments)
7261.235 Million cell updates/sec

Title: US-09-905-114-3
Perfect score: 671
Sequence: 1 acaacaggatctgccccata.....atcatcatgaagaaccatga 671

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	528.6	78.8	890	9	AL546857	AL546857 AL546857
	2	526.6	78.5	896	9	AL546432	AL546432 AL546432
c	3	518.8	77.3	931	9	AL571894	AL571894 AL571894
c	4	511	76.2	833	9	AL572037	AL572037 AL572037
c	5	463	69.0	506	10	AV607154	AV607154 AV607154
	6	458	68.3	666	10	BE048177	BE048177 tz46h09.y
c	7	432.6	64.5	642	10	AW543356	AW543356 C0168H07-
c	8	428.8	63.9	695	13	BI522352	BI522352 603081420
	9	425.4	63.4	825	12	BG867772	BG867772 602786883
	10	422.4	63.0	825	12	BG530136	BG530136 602558626
	11	415.4	61.9	628	13	BI460950	BI460950 603207187
c	12	412.2	61.4	620	10	AW543802	AW543802 C0174C07-
c	13	407.8	60.8	784	9	AI927844	AI927844 wo91g09.x
c	14	407	60.7	610	10	AW540389	AW540389 C0132C04-
	15	406.4	60.6	668	10	AW915564	AW915564 EST346868
	16	402	59.9	832	13	BI838188	BI838188 603083104
c	17	387.6	57.8	634	9	AI660448	AI660448 we67c07.x
	18	387	57.7	685	13	BI833952	BI833952 603085059
	19	377	56.2	759	12	BG534290	BG534290 602553517
c	20	374.2	55.8	566	10	AW538998	AW538998 C0115F01-
	21	363.4	54.2	981	14	BQ710632	BQ710632 AGENCOURT
c	22	352.4	52.5	613	10	AW779563	AW779563 hn81h08.x
	23	351.4	52.4	578	10	AW658627	AW658627 95214 MAR
c	24	351.2	52.3	527	10	AW539589	AW539589 C0121F07-
c	25	335.2	50.0	643	10	AW001760	AW001760 ws04e04.x
	26	329.8	49.2	714	12	BG566893	BG566893 602589586
c	27	321	47.8	542	12	BF434855	BF434855 7o74e07.x
	28	320.8	47.8	855	12	BF237166	BF237166 602028119
c	29	314.2	46.8	554	10	AW243180	AW243180 xn29e10.x
c	30	313	46.6	555	9	AI434430	AI434430 ti14c04.x
	31	309.8	46.2	901	12	BF303508	BF303508 602030047
c	32	309.2	46.1	487	9	AI194886	AI194886 ui58a02.x
c	33	303.8	45.3	497	9	AA564572	AA564572 nj25b03.s
	34	291.2	43.4	586	13	BM489658	BM489658 pgm2n.pk0
	35	284.2	42.4	595	12	BG077563	BG077563 H3016H03-
c	36	283.8	42.3	433	12	BF911384	BF911384 IL2-UT007
	37	279.8	41.7	341	10	AW803005	AW803005 IL2-UM007
c	38	276.6	41.2	511	9	AI271665	AI271665 qj82a02.x
c	39	272.6	40.6	523	9	AI248904	AI248904 qh67e04.x
	40	267.6	39.9	522	14	BQ186387	BQ186387 UI-E-EJ1-
c	41	264.6	39.4	536	9	AA707043	AA707043 zj32d04.s
c	42	263.8	39.3	486	9	AA989198	AA989198 oq95d05.s
	43	262.6	39.1	453	14	H78170	H78170 yu84a11.r1
	44	260.4	38.8	427	10	AW827162	AW827162 xn06e10.y
c	45	251.8	37.5	404	13	BM386137	BM386137 UI-R-CN1-

ALIGNMENTS

RESULT 1
AL546857
LOCUS AL546857 890 bp mRNA linear EST 16-FEB-2001
DEFINITION AL546857 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI026YK08 5 prime, mRNA sequence.
ACCESSION AL546857
VERSION AL546857.1 GI:12880381
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 890)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .890
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI026YK08"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 250 a 215 c 220 g 204 t 1 others
ORIGIN

Query Match 78.8%; Score 528.6; DB 9; Length 890;
Best Local Similarity 87.9%; Pred. No. 4.3e-141;
Matches 576; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          |||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    190 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 249

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          |||||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db    250 CATACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 309

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db    310 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 369

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db    370 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 429

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||||| |||||||||||| |
Db    430 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 489

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db    490 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 549

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    550 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 609

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||||
Db    610 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 669

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db    670 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 729

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| |||||||||||||| | ||
Db    730 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 789

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||||| |   | | |  ||
Db    790 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 844
```

RESULT 2
AL546432
LOCUS AL546432 896 bp mRNA linear EST 16-FEB-2001
DEFINITION AL546432 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI030YD20 5 prime, mRNA sequence.
ACCESSION AL546432
VERSION AL546432.1 GI:12879540
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .896
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI030YD20"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 249 a 217 c 226 g 203 t 1 others
ORIGIN

Query Match 78.5%; Score 526.6; DB 9; Length 896;
Best Local Similarity 87.6%; Pred. No. 1.6e-140;
Matches 574; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

```
Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          |||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    240 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 299

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db    300 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 359

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db    360 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 419

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db    420 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 479

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||||| |||||||||||| |
Db    480 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 539

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db    540 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 599

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    600 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 659

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||||||| ||| |||||||||||||||||||||||||||||||||
Db    660 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 719

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db    720 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 779

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| |||||||||||||| | ||
Db    780 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 839

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||:||||| |   | | |  ||
```

```
Db    840 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTYCAGTTGAATTTAAACTACA 894


RESULT 3
AL571894/c
LOCUS       AL571894                931 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL571894 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI030YD20 3
            prime, mRNA sequence.
ACCESSION   AL571894
VERSION     AL571894.1  GI:12929641
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 931)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .931
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI030YD20"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      209 a    219 c    225 g    271 t      7 others
ORIGIN

  Query Match           77.3%;  Score 518.8;  DB 9;  Length 931;
  Best Local Similarity  86.7%;  Pred. No. 2.9e-138;
  Matches  568;  Conservative    2;  Mismatches   85;  Indels    0;  Gaps    0;

Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          ||||||||:|||||||||||||||||||||||| ||| ||||||||||:| ||||||||
Db    752 ACAACAGGWTCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGRAGAGGCATAA 693

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | ||||||||||||||| ||||||| |||||||||||||||||||||||| |||||||
Db    692 CGTACAACTATGTGATTGGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 633

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||| ||| |||||||| ||     || ||    |||||| ||||||||||
Db    632 TTCTCTACAAGGAAAGGCTGGTGTCTGTGAAGAGGAGTTATNACTACCATGACTATCAGG 573

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db    572 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 513

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||| || ||||||||||  |
Db    512 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAAACATCCGTTAAGG 453

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db    452 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 393

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    392 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 333

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| |||| || ||||||| ||| ||||||||||||||||||||||||||||||||
Db    332 ACATCCGCTTGAGAACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 273

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||||| ||||| || ||| |||||||| ||||||||||||||| | || |
Db    272 CGGTGAAGAAGAGCACCAACTGTGCATATAACAGGATTGTGCTTAGAGGACAAGAAATCG 213

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| |||||||||||||||| | ||
Db    212 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 153

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||| |   | | |  ||
Db    152 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 98


RESULT 4
AL572037/c
LOCUS       AL572037                833 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL572037 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI026YK08 3
            prime, mRNA sequence.
ACCESSION   AL572037
VERSION     AL572037.1  GI:12929919
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 833)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .833
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI026YK08"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      193 a    193 c    191 g    249 t      7 others
ORIGIN

  Query Match           76.2%;  Score 511;  DB 9;  Length 833;
  Best Local Similarity  86.9%;  Pred. No. 4.8e-136;
  Matches  570;  Conservative    2;  Mismatches   83;  Indels    1;  Gaps    1;

Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          |||||||||||||||||||||||||||||||||| :|| |||||||||| | ||||||| |
Db    760 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGWACAGAAATTCAAGGAGAGGCATYA 701

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          |||||||||||||||||||||||| ||||||||||:|||||||||||||| |||||||
Db    700 CATACAACTATGTGATTAGCTCTCGGCTTGGAAGAAAMACATATAAAGAACAATATGCCT 641

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
```

```
Db    640 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 581

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||| |||||||||||||||||||||| |||||||||||||||||||| || ||| |||
Db    580 ATGGANACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 521

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||||| || ||||||||| |
Db    520 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAAACATCCGTTAAGG 461

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db    460 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 401

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    400 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 341

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| |||| || ||||||| ||| ||||||||||||||||||||||||||||||
Db    340 ACATCCGCTTGAGAACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 281

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          |||| |||||||||||| ||||| || |||||||||| |||||||||||||||| | || |
Db    280 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 221

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db    220 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 161

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGAC-CACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||| |||||||||||  |   | | |  ||
Db    160 CTGAAGAGGAGGCCCTGGATGTCAGCGACGCACTTTCCAGTTGAATTTAAACTACA 105
```

RESULT 5
AV607154/c
LOCUS AV607154 506 bp mRNA linear EST 28-NOV-2001
DEFINITION AV607154 Bos taurus kidney fetus Bos taurus cDNA clone E1KI039B05 3', mRNA sequence.
ACCESSION AV607154
VERSION AV607154.1 GI:9737527
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 506)
AUTHORS Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H. and Sugimoto,Y.
TITLE Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs
JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)
MEDLINE 21570554
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
FEATURES Location/Qualifiers
source 1. .506
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="E1KI039B05"
/clone_lib="Bos taurus kidney fetus"
/tissue_type="kidney"
/dev_stage="fetus"
/lab_host="DH10B"
/note="Vector: pZL1; Site_1: Sal1; Site_2: Not1; Poly A was deleted from a Not1 site"
BASE COUNT 124 a 126 c 129 g 126 t 1 others
ORIGIN

Query Match 69.0%; Score 463; DB 10; Length 506;
Best Local Similarity 97.7%; Pred. No. 2.6e-122;
Matches 469; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy    176 TCAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACC 235
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    506 TCAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACC 447

Qy    236 CTACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGT 295
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    446 CTACACCGCTGTCAAGGACTTNGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGT 387

Qy    296 TAGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGA 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    386 TAGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGA 327

Qy    356 GAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTG 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    326 GAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTG 267

Qy    416 GAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGA 475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    266 GAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGA 207

Qy    476 CACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAA 535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    206 CACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAA 147

Qy    536 TATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAG 595
          |||||||||||||| || |||||||||||||||||||||||||||||||||||||||||||
Db    146 TATTGTCAACTCTGTTGTTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAG 87

Qy    596 GTTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||||||||||||||||||||||||||| |||||||||||||| |   | | | |||
Db     86 GTTGTCTGAATCGAAGGCCCTGGATGTCAGTGACCACTTTCCAGTTGAATTTAAACTTCA 27
```

RESULT 6
BE048177
LOCUS BE048177 666 bp mRNA linear EST 20-OCT-2000
DEFINITION tz46h09.y1 NCI_CGAP_Brn52 Homo sapiens cDNA clone IMAGE:2291681 5' similar to SW:DHP2_HUMAN Q13609 DNASE I HOMOLOGOUS PROTEIN DHP2 PRECURSOR ;, mRNA sequence.
ACCESSION BE048177
VERSION BE048177.1 GI:8365155
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 666)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Insert Length: 1048 Std Error: 0.00

```
            Seq primer: -40RP from Gibco
            High quality sequence stop: 431.
FEATURES              Location/Qualifiers
     source           1. .666
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2291681"
                      /clone_lib="NCI_CGAP_Brn52"
                      /tissue_type="tumor, 5 pooled (see description)"
                      /lab_host="DH10B"
                      /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; This library represents the normalized
                      version of NCI_CGAP_Brn35. Cloned unidirectionally.
                      Primer: Oligo dT. Average insert size 1.19 kb. Tumor
                      types include: meningioma, oligodendroglioma, astrocytoma
                      (grade II), medulloblastoma, astrocytoma (grade IV).
                      Constructed by Life Technologies."
BASE COUNT      198 a    149 c    170 g    147 t      2 others
ORIGIN

  Query Match             68.3%;  Score 458;  DB 10;  Length 666;
  Best Local Similarity   88.1%;  Pred. No. 8e-121;
  Matches  497;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
           |||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db      95 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 154

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
           | |||||||||||||||||||||| |||||||||||||||||||||||||||| |||||||
Db     155 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 214

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
           ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db     215 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 274

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
           |||||||||||||||||||||||||| ||||||||||||||||||| || ||| |||
Db     275 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 334

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
           | |||||||| |||||||||||| |||||||||||||||||| ||||||||||||   |
Db     335 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 394

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
           |||| |||||| ||| ||| |||||||| || |||||||   |||||||| || |||||||
Db     395 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 454

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
           ||||||||||||||||||||||| |||||||||||||||||||||||||||||| |||
Db     455 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGNAAGA 514

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
           |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||
Db     515 ACATCCGCTTGAGGACTGACCCCANGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 574

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
           |||| |||||||||| ||||| || |||||||||| ||||||||||||||   | |
Db     575 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAAGAAATC 634

Qy     541 TCAACTCTGGTGGTCCTCAATCAA 564
           ||  ||||     ||  | ||||
Db     635 GCAGTTCTGTGTTCCCCAAGTCAA 658


RESULT 7
AW543356/c
LOCUS       AW543356                642 bp    mRNA    linear   EST 31-AUG-2000
DEFINITION  C0168H07-3 NIA Mouse E7.5 Extraembryonic Portion cDNA Library Mus
            musculus cDNA clone C0168H07 3', mRNA sequence.
ACCESSION   AW543356
VERSION     AW543356.1  GI:7185773
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
            ,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
            III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
  MEDLINE   20381348
COMMENT     Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: C0168  row: H  column: 07
            Seq primer: -21M13 Forward
            High quality sequence stop: 642
            POLYA=Yes.
FEATURES              Location/Qualifiers
     source           1. .642
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="niaEST:C0168H07-3"
                      /db_xref="taxon:10090"
                      /clone="C0168H07"
                      /clone_lib="NIA Mouse E7.5 Extraembryonic Portion cDNA
                      Library"
                      /sex="unknown"
                      /dev_stage="7.5dpc Embryo"
                      /lab_host="DH10B"
                      /note="Vector: pSPORT1 (Gibco/BRL Life Technology);
                      Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
                      5 EPC. The double-stranded cDNA was  synthesized by
                      Gibco's kit with an Oligo(dT) primer [NotI primer-adapter
                      from GibcoBRL]
                      [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3']
                      from 0.8ug of mRNA . The double-stranded cDNAs were
                      treated with T4  DNA polymerase and purified by
                      ethanol-precipitation. The cDNAs were  ligated to
                      Lone-linker LL-Sal3 (include Sal1 sequence). The cDNAs
                      were  purified by phenol/chloroform and separated from
                      free linkers by  Centricon 100. Then, cDNAs were amplified
                      by long-range high fidelity  PCR using Takara's Ex Taq
                      polymerase. Then, the cDNAs were purified by
                      phenol/chloroform and by Centricon 100. The cDNAs were
                      digested with SalI  and NotI enzymes. Then, the cDNAs were
                      size selected by Gibco's Size  Fractionation Column. The
                      cDNAs were cloned into SalI/NotI site of  pSPORT1 plasmid
                      vector. The DH10B E. coli host was transformed with the
                      ligation mixture by chemical method. The library was
                      constructed by  Xiaohong Wang."
BASE COUNT      166 a    161 c    148 g    167 t
ORIGIN

  Query Match             64.5%;  Score 432.6;  DB 10;  Length 642;
  Best Local Similarity   82.6%;  Pred. No. 1.6e-113;
  Matches  495;  Conservative    0;  Mismatches  104;  Indels    0;  Gaps    0;

Qy      57 ATAACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTAT 116
           | |||||||||||||||||||| ||||| |||||||||||||| || ||||| ||||||
Db     642 ACAACATACAACTATGTGATTAGTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTAT 583

Qy     117 GCCTTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTAT 176
           |||||  |||| || || ||||| |||||||| ||   ||  |||| |||||| ||||||
Db     582 GCCTTCGTCTACAAGGAGAAGCTGGTGTCTGTGAAGACAAAATACCACTACCATGACTAT 523

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
           |||| ||||||| |||| |||||||||||||| |||||||||| |||||||| || |||
Db     522 CAGGATGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCC 463
```

```
Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
           |  || |||||||||||||||||||||||||||| ||||||| || || ||||| ||||||
Db     462 TTTACTGCTGTCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTT 403

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
           | |||||| |||||||||| |||||||||| ||||||| | | |  |||||  |||||
Db     402 AAAGAGATAGATGAGCTGGTCGATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAG 343

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
           |||||||| |||||||||| ||||| || ||||| ||||| ||||||||||||||||||
Db     342 AATTTCATCTTCATGGGTGATTTCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGG 283

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
             || |||| ||  |||||||||||||||||| |||||||||| ||||||||||||||||
Db     282 CAGAACATTCGTTTGAGGACGGACCCCAAGTTTGTTTGGCTGATTGGGGACCAAGAGGAC 223

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
           || |||||||||||||| || | ||| |||||||||||||| |||||| | |||||| |
Db     222 ACTACGGTCAAGAAGAGTACCAGCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAG 163

Qy     537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
           || |||||||| |  |  || |  || |    ||||||||| || ||||||||||
Db     162 ATAGTCAACTCCGTGGTCCCCCGTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGAC 103

Qy     597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
           ||||||||   | ||||||||||||||| || |||||||||||  |  |  |  ||
Db     102 TTGTCTGAGGAGGAGGCCCTGGATGTCAGTGATCACTTTCCAGTTGAGTTTAAGCTACA 44
```

RESULT 8
BI522352/c

```
LOCUS       BI522352                695 bp    mRNA    linear   EST 29-AUG-2001
DEFINITION  603081420T1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5220532 3',
            mRNA sequence.
ACCESSION   BI522352
VERSION     BI522352.1  GI:15347144
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 695)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11554  row: h  column: 05
            High quality sequence start: 18
            High quality sequence stop: 695.
FEATURES             Location/Qualifiers
     source          1. .695
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5220532"
                     /clone_lib="NIH_MGC_120"
                     /lab_host="DH10B"
                     /note="Organ: pooled pancreas and spleen; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of spleen and pancreas from 28 yo
                     male.  Library is oligo-dT primed and directionally cloned
                     (EcoRV  site is destroyed upon cloning).  Average insert
                     size 1.5 kb, insert size range 1-2.5 kb.  Library is
                     normalized and enriched for  full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 025. Note: this is a NIH_MGC Library."
```

```
BASE COUNT      159 a    171 c    163 g    202 t
ORIGIN

  Query Match           63.9%;  Score 428.8;  DB 13;  Length 695;
  Best Local Similarity   87.2%;  Pred. No. 2.1e-112;
  Matches  516;  Conservative    0;  Mismatches   72;  Indels    4;  Gaps    4;

Qy      49 GAAAAGGCATAACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAG 108
           | | ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     695 GGAGAGGCATAACATACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAG 636

Qy     109 AACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACC 168
           |||| |||||||||||||| || |||||||| |||||||| ||    || || | |||||
Db     635 CACAATATGCCTTTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACC 576

Qy     169 ACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCC 228
           | |||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||
Db     575 ATGACTATCAGGATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCC 516

Qy     229 AGTCACCCTACA-CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAG 287
           | || ||| ||| | |||||||| |||||||||||| |||||||||||||||||||| |||
Db     515 AATCTCCCCACATCTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAG 456

Qy     288 ACATCCGTTAGAGAGATTGATGAGCTGGCT-GATGTCTACACAGATGTGAAACGTCGCTG 346
           |||||||||  ||||| |||||| ||| | || |||||||| || |||||||  |||||
Db     455 ACATCCGTTAAGGAGATCGATGAGTTGGTTAGAGGTCTACACGGACGTGAAACACCGCTG 396

Qy     347 GAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAA 406
           ||| || |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     395 GAAGGCGGAGAATTTCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAA 336

Qy     407 GAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGA 466
           |||||||||||| |||||||| |||||||| |||||||| ||| ||||||||||||||||||
Db     335 GAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGA 276

Qy     467 CCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAG 526
           |||||||||||||||| |||||||||| ||||| || |||||||||| ||||||||
Db     275 CCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAG 216

Qy     527 AGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCA-AACCTCGTCTTTGATTTCCAGA 585
           ||||||| | || |||| |||| || |||  | ||| |||   || ||||| |||||||
Db     215 AGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAGAACAGTGTTTTTGACTTCCAGA 156

Qy     586 AAGCTTACAGGTTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCC 637
           |||||||| | || |||||  | ||| ||||||||||||||||||| | |
Db     155 AAGCTTACAAGCTGACTGAAGAGGAGG-CCTGGATGTCAGCGACCACTATTC 105
```

RESULT 9
BG867772

```
LOCUS       BG867772                825 bp    mRNA    linear   EST 29-MAY-2001
DEFINITION  602786883F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912893 5',
            mRNA sequence.
ACCESSION   BG867772
VERSION     BG867772.1  GI:14218312
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 825)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
```

```
           Plate: LLAM10817  row: e  column: 22
           High quality sequence stop: 799.
FEATURES             Location/Qualifiers
     source          1. .825
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4912893"
                     /clone_lib="NCI_CGAP_SG2"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
                     NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
                     dT. Average insert size 1.3 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      215 a    203 c    204 g    203 t
ORIGIN

  Query Match           63.4%;  Score 425.4;  DB 12;  Length 825;
  Best Local Similarity   82.7%;  Pred. No. 2.2e-111;
  Matches  522;  Conservative    0;  Mismatches  106;  Indels    3;  Gaps    3;

Qy       2 CAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAAC 61
          |||||  ||||| ||||| ||||||||||||||| || ||||||||| ||| | ||| ||
Db     196 CAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAATTCACGAAGAAGCACAAT 255

Qy      62 ATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCTT 121
           |||||||||||||||||| ||||| |||||||||||||| || ||||| ||||||||||
Db     256 ATACAACTATGTGATTAGTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTATGCCTT 315

Qy     122 TCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGC 181
           ||||  || || ||||| |||||||| ||    ||  |||| |||||| ||||||||||
Db     316 CGTCTACAAGGAGAAGCTGGTGTCTGTGAAGACAAAATACCACTACCATGACTATCAGGA 375

Qy     182 TGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACAC 241
           ||||||| |||| |||||||||||||| |||||||||| |||||||| || ||||  ||
Db     376 TGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTAC 435

Qy     242 CGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGA 301
           ||||||||||||||||||||||||||| ||||||| || || ||||| ||||||| |||
Db     436 TGCTGTCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTTAAAGA 495

Qy     302 GATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTT 361
           ||| |||||||||| |||||||||| ||||||| | | |  |||||  ||||||||||
Db     496 GATAGATGAGCTGGTCGATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTT 555

Qy     362 CATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGA 421
           ||| |||||||||| ||||| || ||||| ||||| ||||||||||||||||||| || |
Db     556 CATCTTCATGGGTGATTTCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAA 615

Qy     422 CATCCGCCTGA-GGACGGACCCCAAGTTCG-TTTGGCTGATCGGGGACCAAGAGGACACC 479
           ||| ||  ||| |||||||||||||||| | |||||||||| ||||||||||||||||||
Db     616 CATTCGTTTGACGGACGGACCCCAAGTTTGTTTTGGCTGATTGGGGACCAAGAGGACACT 675

Qy     480 ACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATT 539
           |||||||||||||| || | ||| |||||||||||||| |||||| | |||||| | ||
Db     676 ACGGTCAAGAAGAGTACCAGCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAGATA 735

Qy     540 GTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTG 599
           |||||||| |  | ||| |  || |    ||||||||| || ||||||||||    |||
Db     736 GTCAACTCCGTGGTTCCCCGTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGACTTG 795

Qy     600 TCTGAATCGAAGGCCCTGGATGTCAGCGACC 630
           |||||  | |  |||||||||||||  || |
Db     796 TCTGACGC-AGAGCCCTGGATGTCACGGATC 825


RESULT 10
BG530136
LOCUS       BG530136                825 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602558626F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4696638 5',
            mRNA sequence.
ACCESSION   BG530136
VERSION     BG530136.1  GI:13521673
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 825)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1524  row: c  column: 07
            High quality sequence stop: 743.
FEATURES             Location/Qualifiers
     source          1. .825
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4696638"
                     /clone_lib="NIH_MGC_61"
                     /tissue_type="embryonal carcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
BASE COUNT      225 a    204 c    217 g    179 t
ORIGIN

  Query Match           63.0%;  Score 422.4;  DB 12;  Length 825;
  Best Local Similarity   83.2%;  Pred. No. 1.6e-110;
  Matches  516;  Conservative    0;  Mismatches  101;  Indels    3;  Gaps    3;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
           |||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db     206 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 265

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
           | |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     266 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 325

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
           ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db     326 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 385

Qy     181 CTGG-AGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTAC 239
           |||  |||||||||||||||||||||| |||||||||||||||||| || ||| ||
Db     386 ATGGCAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCAC 445

Qy     240 ACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGA 299
           || |||||||| |||||||||||| |||||||||||||| ||| ||||||||||||
Db     446 ACTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCAGCCCAGAGACATCCGTTAAG 505

Qy     300 GAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAAT 359
           ||||| |||||| ||| ||| |||||||| || |||||||  |||||||| || | ||||
Db     506 GAGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGCGAAT 565

Qy     360 TTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAA- 418
           |||||||||||||||||||||||||| |||||||||     |||||||||| ||||||
```

```
Db    566 TTCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTTACGTCCCAAGAAGGTCTGGAAC 625

Qy    419 GGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACAC 478
          | ||||||| ||||||| ||||||| ||| | ||||||||||||||||||||||||||||
Db    626 GAACATCCGCTTGAGGACTGACCCCAGGTTTGCTTGGCTGATCGGGGACCAAGAGGACAC 685

Qy    479 CACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGA-TCGTGCTTAGAGGACAAAATA 537
          |||||| |||||||||| ||||| || |||||||||| | |||||||  | || || |
Db    686 CACGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGGTGCTTAGCAGGCACAAGA 745

Qy    538 TTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGT 597
                | |      | |||   || |  |   |||| ||||||||||||| ||||
Db    746 AATCGTCCAGTCTGTCGTCCCAAGTCAAACAGGGTCTTGACTTCCAGAAAGCTTCCAGGC 805

Qy    598 TGTCTGAATCGAAGGCCCTG 617
          || ||| |    ||||| ||
Db    806 TGACTGTAGAAGAGGCCTTG 825


RESULT 11
BI460950
LOCUS       BI460950                628 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603207187F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5272917 5',
            mRNA sequence.
ACCESSION   BI460950
VERSION     BI460950.1  GI:15251606
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 628)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11688  row: n  column: 22
            High quality sequence stop: 628.
FEATURES             Location/Qualifiers
     source          1. .628
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5272917"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      181 a    144 c    166 g    137 t
ORIGIN

  Query Match           61.9%;  Score 415.4;  DB 13;  Length 628;
  Best Local Similarity   90.6%;  Pred. No. 1.5e-108;
  Matches  443;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          |||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    140 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 199

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          |||||||||||||||||||||||| |||||||||||||||||||||||||||| |||||||
Db    200 CATACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 259

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||     || || | |||||| ||||||||||
Db    260 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 319

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
           |||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db    320 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 379

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | |||||||| |||||||||||| |||||||||||||||||| |||||||||||| |
Db    380 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 439

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db    440 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 499

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    500 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 559

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          |||||||| ||||||| ||| |||||||||||||||||||||||||||||||||||||||
Db    560 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 619

Qy    481 CGGTCAAGA 489
          |||| ||||
Db    620 CGGTGAAGA 628


RESULT 12
AW543802/c
LOCUS       AW543802                620 bp    mRNA    linear   EST 31-AUG-2000
DEFINITION  C0174C07-3 NIA Mouse E7.5 Extraembryonic Portion cDNA Library Mus
            musculus cDNA clone C0174C07 3', mRNA sequence.
ACCESSION   AW543802
VERSION     AW543802.1  GI:7186219
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 620)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
            ,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
            III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
  MEDLINE   20381348
COMMENT     Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: C0174  row: C  column: 07
            Seq primer: -21M13 Forward
            High quality sequence stop: 620
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .620
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:C0174C07-3"
                     /db_xref="taxon:10090"
                     /clone="C0174C07"
                     /clone_lib="NIA Mouse E7.5 Extraembryonic Portion cDNA
                     Library"
                     /sex="unknown"
```

/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/note="Vector: pSPORT1 (Gibco/BRL Life Technology); Site_1: SalI; Site_2: NotI; Total RNAs were extracted from 5 EPC. The double-stranded cDNA was synthesized by Gibco's kit with an Oligo(dT) primer [NotI primer-adapter from GibcoBRL]
[5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3']
from 0.8ug of mRNA . The double-stranded cDNAs were treated with T4 DNA polymerase and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal3 (include Sal1 sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang."

BASE COUNT 160 a 159 c 143 g 158 t
ORIGIN

Query Match 61.4%; Score 412.2; DB 10; Length 620;
Best Local Similarity 82.1%; Pred. No. 1.2e-107;
Matches 474; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```
Qy      79 GCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGC 138
           | |||||| |||||||||||||| || ||||| |||||||||||  |||| || || ||||
Db     620 GTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGC 561

Qy     139 TAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGT 198
           | ||||||||| ||    ||  |||| |||||| |||||||||| ||||||| |||| ||||
Db     560 TGGTGTCTGTGAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGT 501

Qy     199 TTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCG 258
           |||||||||| |||||||||| |||||||| || ||||  || ||||||||||||||||
Db     500 TTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCG 441

Qy     259 TGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTG 318
           |||||||||| ||||||| || || ||||| ||||||| |||||| |||||||||  |
Db     440 TGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCG 381

Qy     319 ATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACT 378
           |||||||||| ||||||| | | |  |||||  |||||||||||||| |||||||||| |
Db     380 ATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATT 321

Qy     379 TCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGG 438
           |||| || ||||| ||||| |||||||||||||||||| || |||| ||  |||||||||
Db     320 TCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGAGGACGG 261

Qy     439 ACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAA 498
           |||||||||| |||||||||| |||||||||||||||||| |||||||||||||| || |
Db     260 ACCCCAAGTTTGTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCA 201

Qy     499 ACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTC 558
            ||| |||||||||||||| |||||| | |||||| | || |||||||| |  |  || |
Db     200 GCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTACCCC 141

Qy     559 AATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 618
             || |    ||||||||| || ||||||||||    ||||||||   | |||||||||
Db     140 GTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 81

Qy     619 ATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
           ||||||| || |||||||||||| |   | | |  ||
Db      80 ATGTCAGTGATCACTTTCCAGTTGAGTTTAAGCTACA 44
```

RESULT 13
AI927844/c

```
LOCUS       AI927844                784 bp    mRNA    linear   EST 08-MAR-2000
DEFINITION  wo91g09.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2462752 3'
            similar to SW:DHP2_HUMAN Q13609 DNASE I HOMOLOGOUS PROTEIN DHP2
            PRECURSOR ;, mRNA sequence.
ACCESSION   AI927844
VERSION     AI927844.1  GI:5663808
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 739   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 426.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2462752"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "
BASE COUNT      173 a    188 c    179 g    242 t      2 others
ORIGIN
```

Query Match 60.8%; Score 407.8; DB 9; Length 784;
Best Local Similarity 86.1%; Pred. No. 2.5e-106;
Matches 451; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

```
Qy     132 GAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCA 191
           |||||||| |||||||| ||     || || | |||||| |||||||||| ||||||||||
Db     667 GAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGGATGGAGACGCA 608

Qy     192 GATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAG 251
           |||||||||||||||||| |||| ||||||||||||||| || ||| |||| |||||||||
Db     607 GATGTGTTTTCCAGGGAGCCCTGTGTGGTCTGGTTCCAATCTCCCCACACTGCTGTCAAA 548

Qy     252 GACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAG 311
           |||||||||||  |||||||||||||||||  | ||||||||||||||   ||||| ||||||
Db     547 GACTTCGTGATTATCCCCCTGCACACCACCNCAGAGACATCCGTTAAGGAGATCGATGAG 488

Qy     312 CTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATG 371
           |||  ||| |||||||| || |||||||  ||||||||  || ||||||||||||||||||
Db     487 TTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATTTCATTTTCATG 428

Qy     372 GGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTG 431
           |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||
Db     427 GGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTG 368
```

```
Qy     432 AGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAG 491
           ||||| ||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||
Db     367 AGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTGAAGAAG 308

Qy     492 AGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGT 551
           ||||| ||||| || |||||||||||| |||||||||||||||| | || |||| |||| |
Db     307 AGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTT 248

Qy     552 GGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAG 611
           | |||  | ||||||   || ||||| ||||||||||||||| | || |||||  | ||
Db     247 GTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGACTGAAGAGGAG 188

Qy     612 GCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
           |||||||||||||||||||||||||||||| |   | | |  ||
Db     187 GCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 144
```

RESULT 14
AW540389/c
LOCUS AW540389 610 bp mRNA linear EST 31-AUG-2000
DEFINITION C0132C04-3 NIA Mouse E7.5 Extraembryonic Portion cDNA Library Mus musculus cDNA clone C0132C04 3', mRNA sequence.
ACCESSION AW540389
VERSION AW540389.1 GI:7182806
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 610)
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
TITLE Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
MEDLINE 20381348
COMMENT Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: C0132 row: C column: 04
Seq primer: -21M13 Forward
High quality sequence stop: 610
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .610
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:C0132C04-3"
/db_xref="taxon:10090"
/clone="C0132C04"
/clone_lib="NIA Mouse E7.5 Extraembryonic Portion cDNA Library"
/sex="unknown"
/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/note="Vector: pSPORT1 (Gibco/BRL Life Technology); Site_1: SalI; Site_2: NotI; Total RNAs were extracted from 5 EPC. The double-stranded cDNA was synthesized by Gibco's kit with an Oligo(dT) primer [NotI primer-adapter from GibcoBRL]
[5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3']
from 0.8ug of mRNA . The double-stranded cDNAs were treated with T4 DNA polymerase and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal3 (include Sal1 sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang."
BASE COUNT 157 a 157 c 140 g 156 t
ORIGIN

Query Match 60.7%; Score 407; DB 10; Length 610;
Best Local Similarity 82.4%; Pred. No. 3.8e-106;
Matches 467; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
Qy      89 TGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCTGT 148
           |||||||||||| || ||||| |||||||||||  |||| || || ||||| ||||||||
Db     610 TGGAAGAAACACGTACAAAGAGCAGTATGCCTTCGTCTACAAGGAGAAGCTGGTGTCTGT 551

Qy     149 AAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGGGA 208
           ||   ||  |||| |||||| |||||||||| ||||||| |||| ||||||||||||||
Db     550 GAAGACAAAATACCACTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGGGA 491

Qy     209 ACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTCCC 268
           |||||||||| |||||||| || ||||  || ||||||||||||||||||||||||||||
Db     490 GCCCTTTGTGGTTTGGTTCCATTCCCCCTTTACTGCTGTCAAGGACTTCGTGATTGTCCC 431

Qy     269 CCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTACAC 328
           | ||||||| || || ||||| ||||||| ||||| |||||||||  ||||||||||||
Db     430 CTTGCACACAACTCCCGAGACCTCCGTTAAAGAGATAGATGAGCTGGTCGATGTCTACAC 371

Qy     329 AGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCTGG 388
           ||||||| | | |  |||||  |||||||||||||| |||||||||| ||||| || ||
Db     370 GGATGTGAGAAGCCAGTGGAAGACAGAGAATTTCATCTTCATGGGTGATTTCAACGCCGG 311

Qy     389 CTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCCAAGTT 448
           ||| ||||| ||||||||||||||||| || |||| ||  ||||||||||||||||||||
Db     310 CTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAACATTCGTTTGAGGACGGACCCCAAGTT 251

Qy     449 CGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCCTA 508
           |||||||||| |||||||||||||||||| |||||||||||||| || | ||| |||||
Db     250 TGTTTGGCTGATTGGGGACCAAGAGGACACTACGGTCAAGAAGAGTACCAGCTGTGCCTA 191

Qy     509 TGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAACCT 568
           ||||||||| |||||| | |||||| | || |||||||| |  | ||| |  || |
Db     190 TGACAGGATTGTGCTTTGTGGACAAGAGATAGTCAACTCCGTGGTTCCCCGTTCCAGTGG 131

Qy     569 CGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGGATGTCAGCGA 628
           ||||||||| || |||||||||||    |||||||   | ||||||||||||||| ||
Db     130 CGTCTTTGACTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGGATGTCAGTGA 71

Qy     629 CCACTTTCCAGTTCATCATCATCATCA 655
           |||||||||||| |   | | |  ||
Db      70 TCACTTTCCAGTTGAGTTTAAGCTACA 44
```

RESULT 15
AW915564
LOCUS AW915564 668 bp mRNA linear EST 25-MAY-2000
DEFINITION EST346868 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone RGICW80 5' end, mRNA sequence.
ACCESSION AW915564
VERSION AW915564.1 GI:8081263
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 668)
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.
TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat Gene Index
JOURNAL Unpublished (1998)

```
COMMENT     Contact: Lee, NH
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Tel: (301)-838-3529
            Fax: (301)-838-0208
            Email: nhlee@tigr.org
            This clone is available through the ATCC, contact the ATCC
            tel#703-365-2700 for further information
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="Rattus sp."
                     /db_xref="taxon:10118"
                     /clone="RGICW80"
                     /clone_lib="Normalized rat embryo, Bento Soares"
                     /dev_stage="embryo 8, 12, 18 dpc"
                     /note="Vector: pT7T3Pac; Site_1: EcoRI; Site_2: NotI"
BASE COUNT      172 a    160 c    168 g    168 t
ORIGIN

  Query Match             60.6%;  Score 406.4;  DB 10;  Length 668;
  Best Local Similarity   84.9%;  Pred. No. 5.8e-106;
  Matches  467;  Conservative    0;  Mismatches   81;  Indels    2;  Gaps    1;

Qy     106 AAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCT 165
           |||||||||||||||| ||||| || || ||||| |||||||| ||   ||   ||||||
Db       2 AAGAACAGTATGCCTTCCTCTACAAGGAGAAGCTGGTGTCTGTGAAGGCAAA--ACCTCT 59

Qy     166 ACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGT 225
           |||| |||||||||| ||||||| |||| |||||||||||||| |||||||||| ||||
Db      60 ACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGT 119

Qy     226 TCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTG 285
           ||||| | |||| ||| |||| |||||||||||||||||||||| ||||||| || ||||
Db     120 TCCAGGCGCCCTTCACTGCTGCCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCTG 179

Qy     286 AGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCT 345
           | || ||||||| |||||| |||||||||||||| |||||||| ||||||| | | || |
Db     180 AAACCTCCGTTAAAGAGATAGATGAGCTGGCTGACGTCTACACGGATGTGAGAAGACGAT 239

Qy     346 GGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCA 405
           |||| |||||||||||||| |||||||||| ||||||||||||||||||||||||||||
Db     240 GGAAGGCAGAGAATTTCATCTTCATGGGTGATTTCAATGCTGGCTGCAGCTACGTCCCCA 299

Qy     406 AGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGG 465
           |||||||||||||| ||||||| ||||||| |||||||| || |||||||||| ||||
Db     300 AGAAGGCCTGGAAGAACATCCGTTTGAGGACAGACCCCAACTTTGTTTGGCTGATTGGGG 359

Qy     466 ACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTA 525
           |||||||||||||||||||||||||||||| | ||| |||||||||||||| ||||||
Db     360 ACCAAGAGGACACCACGGTCAAGAAGAGCACCAGCTGTGCCTATGACAGGATTGTGCTTC 419

Qy     526 GAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGA 585
           | |||||| | || |||||||||| | ||| | || |    ||||||||| || ||||
Db     420 GCGGACAAGAGATAGTCAACTCTGTGGTTCCCCGCTCCAGTGGCGTCTTTGACTTTCAGA 479

Qy     586 AAGCTTACAGGTTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATC 645
           |||||||   |||||||||| | ||||||||||||||| ||||||||||||| |
Db     480 AAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGGATGTCAGTGACCACTTTCCAGTTGAGT 539

Qy     646 ATCATCATCA 655
           | | |  ||
Db     540 TTAAGCTACA 549


Search completed: December  9, 2002, 12:18:12
Job time : 1505.6 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 08:59:53 ; Search time 40.9082 Seconds
(without alignments)
5030.295 Million cell updates/sec

Title: US-09-905-114-3
Perfect score: 671
Sequence: 1 acaacaggatctgccccata.....atcatcatgaagaaccatga 671

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

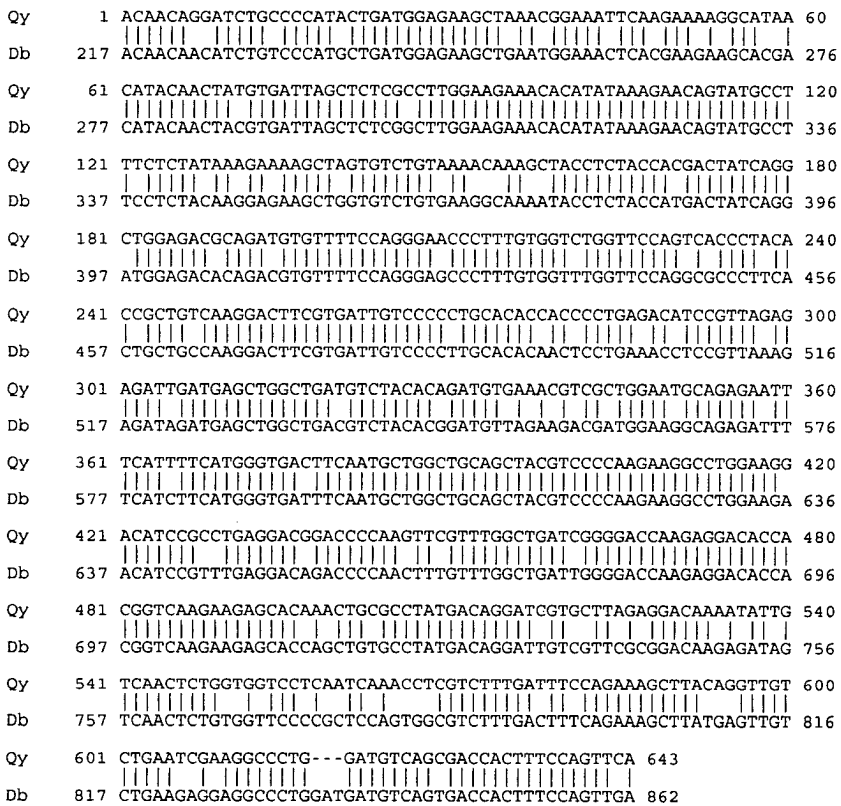
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	481.2	71.7	1208	1	US-08-640-765A-2	Sequence 2, Appli
	2	481.2	71.7	1208	3	US-09-073-613-2	Sequence 2, Appli
	3	163.2	24.3	1039	1	US-07-895-300A-2	Sequence 2, Appli
	4	163.2	24.3	1039	4	US-09-638-112-2	Sequence 2, Appli
	5	163.2	24.3	1039	5	PCT-US93-05136-2	Sequence 2, Appli
	6	163.2	24.3	1039	5	PCT-US93-10519-1	Sequence 1, Appli
	7	161.6	24.1	1039	1	US-08-458-367-2	Sequence 2, Appli
	8	124.6	18.6	1615	2	US-08-491-988-15	Sequence 15, Appl
	9	124.6	18.6	1624	2	US-08-491-988-14	Sequence 14, Appl
	10	124.6	18.6	1648	2	US-08-491-988-13	Sequence 13, Appl
	11	119	17.7	1055	2	US-08-468-012A-1	Sequence 1, Appli
	12	119	17.7	1055	4	US-09-054-989-1	Sequence 1, Appli
	13	35.2	5.2	1427	3	US-09-173-581-10	Sequence 10, Appl
	14	35.2	5.2	1427	4	US-09-420-915-10	Sequence 10, Appl
	15	32.8	4.9	1293	2	US-08-924-440-1	Sequence 1, Appli
	16	32.2	4.8	289	4	US-09-007-005-17	Sequence 17, Appl
	17	32.2	4.8	289	4	US-09-244-796-17	Sequence 17, Appl
c	18	32.2	4.8	7218	1	US-08-232-463-14	Sequence 14, Appl
	19	31.8	4.7	1511	2	US-08-809-763-3	Sequence 3, Appli
	20	31.8	4.7	1511	3	US-08-956-253-3	Sequence 3, Appli
c	21	31.6	4.7	1001	4	US-09-641-638-90	Sequence 90, Appl
c	22	31.2	4.6	248	3	US-08-513-974B-300	Sequence 300, App
c	23	31.2	4.6	248	3	US-08-513-974B-301	Sequence 301, App
	24	30.6	4.6	1920	1	US-08-487-753-1	Sequence 1, Appli
	25	30.6	4.6	1920	2	US-08-480-065-1	Sequence 1, Appli
	26	30.6	4.6	1920	3	US-08-487-744-1	Sequence 1, Appli
	27	30.6	4.6	1920	5	PCT-US93-09167-1	Sequence 1, Appli
c	28	30.6	4.6	2900	2	US-09-027-337-9	Sequence 9, Appli
c	29	30.6	4.6	2900	4	US-09-644-600-9	Sequence 9, Appli
c	30	30.6	4.6	3147	2	US-09-027-337-1	Sequence 1, Appli
c	31	30.6	4.6	3147	4	US-09-644-600-1	Sequence 1, Appli
	32	30.6	4.6	3147	4	US-09-644-600-18	Sequence 18, Appl
c	33	30	4.5	1734	4	US-09-484-970B-63	Sequence 63, Appl
	34	30	4.5	2071	1	US-07-923-724-1	Sequence 1, Appli
	35	30	4.5	2071	2	US-08-609-426A-1	Sequence 1, Appli
	36	30	4.5	2071	2	US-08-374-652C-3	Sequence 3, Appli
c	37	30	4.5	4049	1	US-08-162-809-17	Sequence 17, Appl
c	38	30	4.5	4097	1	US-08-162-809-11	Sequence 11, Appl
	39	29.8	4.4	3981	2	US-08-955-138-2	Sequence 2, Appli
c	40	29.6	4.4	47	4	US-09-146-054-7	Sequence 7, Appli
c	41	29.6	4.4	47	4	US-09-664-977A-7	Sequence 7, Appli
	42	29.6	4.4	270	4	US-09-146-054-8	Sequence 8, Appli
	43	29.6	4.4	270	4	US-09-664-977A-8	Sequence 8, Appli
c	44	29.6	4.4	352	4	US-09-228-986-21	Sequence 21, Appl
c	45	29.6	4.4	560	3	US-09-042-426-5	Sequence 5, Appli

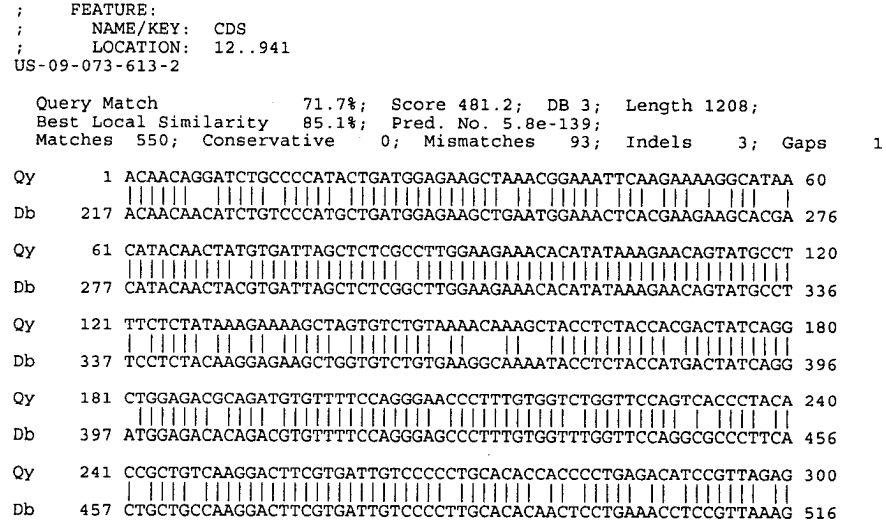
ALIGNMENTS

```
RESULT 1
US-08-640-765A-2
; Sequence 2, Application US/08640765A
; Patent No. 5821103
;  GENERAL INFORMATION:
;    APPLICANT:  Tanuma, Sei-ichi
;    TITLE OF INVENTION:  NOVEL DEOXYRIBONUCLEASE
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/640,765A
;      FILING DATE:  06-MAY-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6239518
;      FILING DATE:  06-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Green, Robert F.
;      REGISTRATION NUMBER:  27555
;      REFERENCE/DOCKET NUMBER:  73362
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5600
;      TELEFAX:  (312) 616-5700
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1208 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  12..941
US-08-640-765A-2

  Query Match             71.7%;  Score 481.2;  DB 1;  Length 1208;
  Best Local Similarity   85.1%;  Pred. No. 5.8e-139;
  Matches  550;  Conservative    0;  Mismatches   93;  Indels    3;  Gaps    1;
```



RESULT 2
US-09-073-613-2
; Sequence 2, Application US/09073613
; Patent No. 6143875
; GENERAL INFORMATION:
; APPLICANT: Tanuma, Sei-ichi
; TITLE OF INVENTION: ANTIBODY TO NOVEL DEOXYRIBONUCLEASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,613
; FILING DATE: 06-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/640,765
; FILING DATE:
; APPLICATION NUMBER: JP 6239518
; FILING DATE: 06-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: LARCHER, Carol
; REGISTRATION NUMBER: 35243
; REFERENCE/DOCKET NUMBER: 85519
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1208 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12..941
US-09-073-613-2

Query Match 71.7%; Score 481.2; DB 3; Length 1208;
Best Local Similarity 85.1%; Pred. No. 5.8e-139;
Matches 550; Conservative 0; Mismatches 93; Indels 3; Gaps 1;



RESULT 3
US-07-895-300A-2
; Sequence 2, Application US/07895300A

```
; Patent No. 5279823
;  GENERAL INFORMATION:
;    APPLICANT:  Frenz, John
;    APPLICANT:  Shire, Steven J.
;    APPLICANT:  Sliwkowski, Mary B.
;    TITLE OF INVENTION:  PURIFIED FORMS OF DNase
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/895,300A
;      FILING DATE:  19920608
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hensley, Max D.
;      REGISTRATION NUMBER:  27,043
;      REFERENCE/DOCKET NUMBER:  747
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1994
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1039 bases
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-895-300A-2

  Query Match             24.3%;  Score 163.2;  DB 1;  Length 1039;
  Best Local Similarity   57.4%;  Pred. No. 9.5e-41;
  Matches  335;  Conservative    0;  Mismatches  243;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 176
           ||  | || |      | |  |||||||      | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
           || |  | |     ||    ||   | | || ||      | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
           | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |    | ||
Db     589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
           |||||  || |  ||   ||| ||||||   |||||  ||       |||      |||
Db     649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
           |  |||| || ||||| ||||||||||| ||||||||||| ||              |||
Db     709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
           |||||||||| ||||  |||||  |||  |||||||||   |||   |  |||
Db     769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
           |||||  |     |   |||  |||| ||||||||||||||||||  ||  |||
Db     829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy     537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
           |      | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db     886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy     597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
           ||  |||      |||||  |   |||| |||||| ||||||
Db     946 CTGAGTGACCAACTGGCCCAAGCCATCAGTGACCACTATCCAGT 989


RESULT 4
US-09-638-112-2
; Sequence 2, Application US/09638112
; Patent No. 6440412
;  GENERAL INFORMATION:
;    APPLICANT:  Frenz, John
;    APPLICANT:  Shire, Steven J.
;    APPLICANT:  Sliwkowski, Mary B.
;    TITLE OF INVENTION:  PURIFIED FORMS OF DNase
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  1 DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/638,112
;      FILING DATE:  09-Aug-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/942561
;      FILING DATE:  01-OCT-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Johnston, Sean A.
;      REGISTRATION NUMBER:  35,910
;      REFERENCE/DOCKET NUMBER:  P0747C7
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650/225-3562
;      TELEFAX:  650/952-9881
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1039 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-09-638-112-2

  Query Match             24.3%;  Score 163.2;  DB 4;  Length 1039;
  Best Local Similarity   57.4%;  Pred. No. 9.5e-41;
  Matches  335;  Conservative    0;  Mismatches  243;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 176
           ||  | || |      | |  |||||||      | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
           || |  | |     ||    ||   | | || ||      | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588
```

```
Qy    237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
          | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |    | ||
Db    589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy    297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
            ||||| || |  ||   ||| ||||||   |||||  ||       |||      |||
Db    649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy    357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
          |  |||| || ||||| ||||||||||| ||                  |||
Db    709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy    417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
            ||||||||||| ||||  |||||  |||  |||||||||  |||  |  |||
Db    769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy    477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
          |||||  |      |    |||  |||| ||||||||||||||||||  ||  |||
Db    829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy    537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
            |       | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db    886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy    597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
            ||  |||       |||||  |   |||| |||||| ||||||
Db    946 CTGAGTGACCAACTGGCCCAAGCCATCAGTGACCACTATCCAGT 989
```

RESULT 5
PCT-US93-05136-2
; Sequence 2, Application PC/TUS9305136
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: PURIFIED FORMS OF DNase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05136
; FILING DATE: 19930528
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: P35,910
; REFERENCE/DOCKET NUMBER: 747PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3562
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1039 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US93-05136-2

Query Match 24.3%; Score 163.2; DB 5; Length 1039;

Best Local Similarity 57.4%; Pred. No. 9.5e-41;
Matches 335; Conservative 0; Mismatches 243; Indels 6; Gaps 2;

```
Qy     60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
          || ||  |||| ||| | ||    |  || ||| | ||||  |||||  || |  ||
Db    409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy    120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 176
          ||   | || |      | |  |||||||      | ||||||  ||||  | ||   ||
Db    469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy    177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
          ||  | | |      ||   ||   | | || ||      | ||| |||||   || |
Db    529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy    237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
          | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |    | ||
Db    589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy    297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
            ||||| || |  ||   ||| ||||||   |||||  ||       |||      |||
Db    649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy    357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
          |  |||| || ||||| ||||||||||| ||||||||||  ||                |||
Db    709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy    417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
            |||||||||| ||||  |||||  |||  |||||||||  |||  |  |||
Db    769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy    477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
          |||||  |      |    |||  |||| ||||||||||||||||||  ||  |||
Db    829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy    537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
            |       | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db    886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy    597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
            ||  |||       |||||  |   |||| |||||| ||||||
Db    946 CTGAGTGACCAACTGGCCCAAGCCATCAGTGACCACTATCCAGT 989
```

RESULT 6
PCT-US93-10519-1
; Sequence 1, Application PC/TUS9310519
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Compaction Assay for Assessment of Respiratory Disease
; TITLE OF INVENTION: Therapy
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10519
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.

```
;      REGISTRATION NUMBER:  35,910
;      REFERENCE/DOCKET NUMBER:  792C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3562
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1039 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US93-10519-1

  Query Match             24.3%;  Score 163.2;  DB 5;  Length 1039;
  Best Local Similarity   57.4%;  Pred. No. 9.5e-41;
  Matches  335;  Conservative    0;  Mismatches  243;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 176
           ||  | || |       | |  |||||||       | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
           || |  | |      ||    ||   | | || ||      | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
           | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |    | ||
Db     589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
             ||||| || |  ||   ||| ||||||  |||||  ||      |||      |||
Db     649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
           |  |||| || ||||| |||||||||| |||||||||| ||              |||
Db     709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
           |||||||||| ||||  |||||  |||  |||||||||  |||  |  |||
Db     769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
           |||||  |     |   |||  |||| ||||||||||||||||||  ||  |||
Db     829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy     537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
           |      | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db     886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy     597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
           ||  |||      |||||  |   |||| ||||||| ||||||
Db     946 CTGAGTGACCAACTGGCCCAAGCCATCAGTGACCACTATCCAGT 989


RESULT 7
US-08-458-367-2
; Sequence 2, Application US/08458367
; Patent No. 5783433
;  GENERAL INFORMATION:
;    APPLICANT:  Frenz, John
;    APPLICANT:  Shire, Steven J.
;    APPLICANT:  Sliwkowski, Mary B.
;    TITLE OF INVENTION:  PURIFIED FORMS OF DNase
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,367
;      FILING DATE:  02-Jun-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/409631
;      FILING DATE:  22-Mar-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/348284
;      FILING DATE:  30-No. 5783433-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/116186
;      FILING DATE:  02-Sep-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/895300
;      FILING DATE:  08-Jun-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Johnston, Sean A.
;      REGISTRATION NUMBER:  35,910
;      REFERENCE/DOCKET NUMBER:  P0747C4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3562
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1039 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-08-458-367-2

  Query Match             24.1%;  Score 161.6;  DB 1;  Length 1039;
  Best Local Similarity   57.2%;  Pred. No. 3e-40;
  Matches  334;  Conservative    0;  Mismatches  244;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 176
           ||  | || |       | |  |||||||       | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
           || |  | |      ||    ||   | | || ||      | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
           | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |      ||
Db     589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACCGAGTA 648

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
             ||||| || |  ||   ||| ||||||  |||||  ||      |||      |||
Db     649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
           |  |||| || ||||| |||||||||| |||||||||| ||              |||
Db     709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
           |||||||||| ||||  |||||  |||  |||||||||  |||  |  |||
Db     769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
```

```
Db    829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy    537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596

Db    886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy    597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640

Db    946 CTGAGTGACCAACTGGCCCAAGCCATCAGTGACCACTATCCAGT 989


RESULT 8
US-08-491-988-15
; Sequence 15, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1615 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-15

  Query Match             18.6%;  Score 124.6;  DB 2;  Length 1615;
  Best Local Similarity   53.2%;  Pred. No. 1e-28;
  Matches  312;  Conservative    0;  Mismatches  269;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119

Db    1009 ACCTACCACTACGTTGTTTCTGAACCCCTCGGGCGTAACTCTTACAAAGAACGGTACCTG 1068

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGAC---TAT 176

Db    1069 TTCCTGTTCCGTCCGAACAAAGTTTCAGTACTGGATACCTACCAGTACGACGACGGATGC 1128

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236

Db    1129 GAATCTTGCGGTAACGACTCTTTCTCCCGGGAACCGGCTGTTGTTAAATTCTCGAGCCAC 1188

Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296

Db    1189 TCTACCAAGGTTAAAGAGTTCGCTATCGTTGCTCTGCACAGCGCGCCGTCTGACGCTGTT 1248

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356

Db    1249 GCTGAAATCAACTCTCTGTACGACGTTTACCTGGACGTTCAGCAGAAATGGCACCTGAAC 1308

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416

Db    1309 GACGTCATGCTGATGGGTGACTTCAACGCTGACTGCTCTTATGTAACCTCTTCTCAGTGG 1368

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476

Db    1369 TCATCGATTCGTCTGCGCACCTCGTCGACCTTCCAGTGGCTGATCCCGGACTCCGCTGAC 1428

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536

Db    1429 ACCACCGCTACT---AGTACCAACTGCGCTTACGACCGTATCGTTGTTGCTGGATCCCTG 1485

Qy     537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596

Db    1486 CTGCAGTCTTCTGTTGTACCGGGTAGCGCGGCCCCGTTCGACTTCCAGGCTGCATATGGT 1545

Qy     597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCA 643

Db    1546 CTTTCGAACGAAATGGCGCTGGCCATCTCTGATCACTACCCGGTTGA 1592


RESULT 9
US-08-491-988-14
; Sequence 14, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1624 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-14

  Query Match             18.6%;  Score 124.6;  DB 2;  Length 1624;
  Best Local Similarity   53.2%;  Pred. No. 1e-28;
  Matches  312;  Conservative    0;  Mismatches  269;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
```

```
Db   1018 ACCTACCACTACGTTGTTTCTGAACCCCTCGGGCGTAACTCTTACAAAGAACGGTACCTG 1077

Qy    120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGAC---TAT 176

Db   1078 TTCCTGTTCCGTCCGAACAAAGTTTCAGTACTGGATACCTACCAGTACGACGACGGATGC 1137

Qy    177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236

Db   1138 GAATCTTGCGGTAACGACTCTTTCTCCCGGGAACCGGCTGTTGTTAAATTCTCGAGCCAC 1197

Qy    237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296

Db   1198 TCTACCAAGGTTAAAGAGTTCGCTATCGTTGCTCTGCACAGCGCGCCGTCTGACGCTGTT 1257

Qy    297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356

Db   1258 GCTGAAATCAACTCTCTGTACGACGTTTACCTGGACGTTCAGCAGAAATGGCACCTGAAC 1317

Qy    357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416

Db   1318 GACGTCATGCTGATGGGTGACTTCAACGCTGACTGCTCTTATGTAACCTCTTCTCAGTGG 1377

Qy    417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476

Db   1378 TCATCGATTCGTCTGCGCACCTCGTCGACCTTCCAGTGGCTGATCCCGGACTCCGCTGAC 1437

Qy    477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536

Db   1438 ACCACCGCTACT---AGTACCAACTGCGCTTACGACCGTATCGTTGTTGCTGGATCCCTG 1494

Qy    537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596

Db   1495 CTGCAGTCTTCTGTTGTACCGGGTAGCGCGGCCCCGTTCGACTTCCAGGCTGCATATGGT 1554

Qy    597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCA 643

Db   1555 CTTTCGAACGAAATGGCGCTGGCCATCTCTGATCACTACCCGGTTGA 1601


RESULT 10
US-08-491-988-13
; Sequence 13, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1648 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-13

  Query Match            18.6%;  Score 124.6;  DB 2;  Length 1648;
  Best Local Similarity   53.2%;  Pred. No. 1e-28;
  Matches  312;  Conservative    0;  Mismatches  269;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119

Db    1042 ACCTACCACTACGTTGTTTCTGAACCCCTCGGGCGTAACTCTTACAAAGAACGGTACCTG 1101

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGAC---TAT 176

Db    1102 TTCCTGTTCCGTCCGAACAAAGTTTCAGTACTGGATACCTACCAGTACGACGACGGATGC 1161

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236

Db    1162 GAATCTTGCGGTAACGACTCTTTCTCCCGGGAACCGGCTGTTGTTAAATTCTCGAGCCAC 1221

Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296

Db    1222 TCTACCAAGGTTAAAGAGTTCGCTATCGTTGCTCTGCACAGCGCGCCGTCTGACGCTGTT 1281

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356

Db    1282 GCTGAAATCAACTCTCTGTACGACGTTTACCTGGACGTTCAGCAGAAATGGCACCTGAAC 1341

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416

Db    1342 GACGTCATGCTGATGGGTGACTTCAACGCTGACTGCTCTTATGTAACCTCTTCTCAGTGG 1401

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476

Db    1402 TCATCGATTCGTCTGCGCACCTCGTCGACCTTCCAGTGGCTGATCCCGGACTCCGCTGAC 1461

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536

Db    1462 ACCACCGCTACT---AGTACCAACTGCGCTTACGACCGTATCGTTGTTGCTGGATCCCTG 1518

Qy     537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596

Db    1519 CTGCAGTCTTCTGTTGTACCGGGTAGCGCGGCCCCGTTCGACTTCCAGGCTGCATATGGT 1578

Qy     597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCA 643

Db    1579 CTTTCGAACGAAATGGCGCTGGCCATCTCTGATCACTACCCGGTTGA 1625


RESULT 11
US-08-468-012A-1
; Sequence 1, Application US/08468012A
; Patent No. 5830744
;  GENERAL INFORMATION:
;    APPLICANT:  RUBEN, ET AL.
;    TITLE OF INVENTION:  Human DNase
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
```

```
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WORD PERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/468,012A
;       FILING DATE:  6 JUNE 1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/04954
;       FILING DATE:  5 MAY 1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  MULLINS, J.G.
;       REGISTRATION NUMBER:  33,073
;       REFERENCE/DOCKET NUMBER:  325800-428
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-994-1700
;       TELEFAX:  201-994-1744
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1055 BASE PAIRS
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
;     MOLECULE TYPE:  cDNA
US-08-468-012A-1

  Query Match             17.7%;  Score 119;  DB 2;  Length 1055;
  Best Local Similarity   55.4%;  Pred. No. 4.3e-27;
  Matches  252;  Conservative    1;  Mismatches  196;  Indels    6;  Gaps    1;

Qy     186 GACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCT 245
           || |  || || ||| || |||| || ||||||| |  ||||    |  |||  ||  |
Db     359 GAGGATGACGTCTTTGCCCGGGAGCCATTTGTGGCCCAGTTCTCTTTGCCCAGCAATGTC 418

Qy     246 GTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATT 305
           |      | | ||| | ||||| |||||||||| ||| || |     |  | ||| |
Db     419 CTTCCCAGCCTGGTGTTGGTCCCGCTGCACACCACTCCTAAGGCCGTAGAGAAGGAGCTG 478

Qy     306 GATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATT 365
           | |  ||    ||||| |     || ||    |  | |||| |     || |  | ||
Db     479 AACGCCCTCTACGATGTGTTTCTGGAGGTCTCCCAGCACTGGCAGAGCAAGGACGTGATC 538

Qy     366 TTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATC 425
           |  | || |||||||||||||| ||||    |    |  |||| |||  |  | |  |  |
Db     539 CTGCTTGGGGACTTCAATGCTGACTGCGCTTCACTGACCAAAAAGCGCCTGGACAAGCTG 598

Qy     426 CGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTC 485
           : ||| |||| || ||    |||   ||| |||| |  ||    |||||||||| ||
Db     599 GSGCTGCGGACTGAGCCAGGCTTCCACTGGGTGATTGCCGATGGGGAGGACACCACAGTG 658

Qy     486 AAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAAC 545
           |   |||||  ||||| |||||||| |  |||||||    ||  |          | |
Db     659 CGGGCCAGCACCCACTGCACCTATGACCGCGTCGTGCTGCACGGGGA------GCGCTGC 712

Qy     546 TCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAA 605
           ||   |  ||  |       | |||||| ||||  |     | |  |  |  | ||
Db     713 CGGAGTCTGCTGCACACTGCGGCTGCCTTTGACTTCCCCACGAGCTTCCAGCTCACCGAG 772

Qy     606 TCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
           | |||||||  |  |||| |||||||  || ||
Db     773 GAGGAGGCCCTCAACATCAGTGACCACTACCCCGT 807


RESULT 12
US-09-054-989-1
; Sequence 1, Application US/09054989
; Patent No. 6251648
;  GENERAL INFORMATION:
;    APPLICANT:  RUBEN, ET AL.
;    TITLE OF INVENTION:  Human DNase
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/054,989
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/468,012
;      FILING DATE:  6-JUNE-1995
;      APPLICATION NUMBER:  PCT/US94/04954
;      FILING DATE:  5 MAY 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MULLINS, J.G.
;      REGISTRATION NUMBER:  33,073
;      REFERENCE/DOCKET NUMBER:  325800-428
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1055 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  cDNA
US-09-054-989-1

  Query Match             17.7%;  Score 119;  DB 4;  Length 1055;
  Best Local Similarity   55.4%;  Pred. No. 4.3e-27;
  Matches  252;  Conservative    1;  Mismatches  196;  Indels    6;  Gaps    1;

Qy     186 GACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCT 245
           || |  || || ||| || |||| || ||||||| |  ||||    |  |||  ||  |
Db     359 GAGGATGACGTCTTTGCCCGGGAGCCATTTGTGGCCCAGTTCTCTTTGCCCAGCAATGTC 418

Qy     246 GTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATT 305
           |      | | ||| | ||||| |||||||||| ||| || |     |  | ||| |
Db     419 CTTCCCAGCCTGGTGTTGGTCCCGCTGCACACCACTCCTAAGGCCGTAGAGAAGGAGCTG 478

Qy     306 GATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATT 365
           | |  ||    ||||| |     || ||    |  | |||| |     || |  | ||
Db     479 AACGCCCTCTACGATGTGTTTCTGGAGGTCTCCCAGCACTGGCAGAGCAAGGACGTGATC 538

Qy     366 TTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATC 425
           |  | || |||||||||||||| ||||    |    |  |||| |||  |  | |  |  |
Db     539 CTGCTTGGGGACTTCAATGCTGACTGCGCTTCACTGACCAAAAAGCGCCTGGACAAGCTG 598

Qy     426 CGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTC 485
           : ||| |||| || ||    |||   ||| |||| |  ||    |||||||||| ||
Db     599 GSGCTGCGGACTGAGCCAGGCTTCCACTGGGTGATTGCCGATGGGGAGGACACCACAGTG 658

Qy     486 AAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAAC 545
           |   |||||  ||||| |||||||| |  |||||||    ||  |          | |
Db     659 CGGGCCAGCACCCACTGCACCTATGACCGCGTCGTGCTGCACGGGGA------GCGCTGC 712

Qy     546 TCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAA 605
           ||   |  ||  |       | |||||| ||||  |     | |  |  |  | ||
Db     713 CGGAGTCTGCTGCACACTGCGGCTGCCTTTGACTTCCCCACGAGCTTCCAGCTCACCGAG 772

Qy     606 TCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
           | |||||||  |  |||| |||||||  || ||
Db     773 GAGGAGGCCCTCAACATCAGTGACCACTACCCCGT 807


RESULT 13
```

```
US-09-173-581-10
; Sequence 10, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Yue, Henry
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Gorgone, Gina
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Lu, Aina
;  TITLE OF INVENTION: Protein Kinase Homologs
;  FILE REFERENCE: PF-0614 US
;  CURRENT APPLICATION NUMBER: US/09/173,581A
;  CURRENT FILING DATE: 1998-10-15
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PERL Program
; SEQ ID NO 10
;   LENGTH: 1427
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: 119819
US-09-173-581-10

  Query Match             5.2%;  Score 35.2;  DB 3;  Length 1427;
  Best Local Similarity   53.7%;  Pred. No. 0.39;
  Matches   73;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy     348 AATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAG 407
           ||||||||  ||  |||   ||| |   |     |   |     |  | |||  ||  | ||||
Db     857 AATGCAGATAAGATCAAGATCAAGATCGCAGACCTGGGCAACGCCTGCTGGGTGCACAAG 916

Qy     408 AAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGAC 467
           |   |  | | |||||||       ||  | |||       |||    |||||||||  | |
Db     917 CACTTCACGGAAGACATCCAGACTCGGCAGTACCGGGCCGTCGAGGTGCTGATCGGCGCC 976

Qy     468 CAAGAGGACACCACGG 483
           || | | | || |||
Db     977 GAATACGGCCCCCCGG 992


RESULT 14
US-09-420-915-10
; Sequence 10, Application US/09420915
; Patent No. 6264947
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Yue, Henry
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Gorgone, Gina
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Lu, Aina
;  TITLE OF INVENTION: Protein Kinase Homologs
;  FILE REFERENCE: PF-0614 US
;  CURRENT APPLICATION NUMBER: US/09/420,915
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/173,581
;  EARLIER FILING DATE: 1998-10-15
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PERL Program
; SEQ ID NO 10
;   LENGTH: 1427
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: 119819
US-09-420-915-10

  Query Match             5.2%;  Score 35.2;  DB 4;  Length 1427;
  Best Local Similarity   53.7%;  Pred. No. 0.39;
  Matches   73;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy     348 AATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAG 407
           ||||||||  ||  |||   ||| |   |     |   |     |  | |||  ||  | ||||
Db     857 AATGCAGATAAGATCAAGATCAAGATCGCAGACCTGGGCAACGCCTGCTGGGTGCACAAG 916

Qy     408 AAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGAC 467
           |   |  | | |||||||       ||  | |||       |||    |||||||||  | |
Db     917 CACTTCACGGAAGACATCCAGACTCGGCAGTACCGGGCCGTCGAGGTGCTGATCGGCGCC 976

Qy     468 CAAGAGGACACCACGG 483
           || | | | || |||
Db     977 GAATACGGCCCCCCGG 992


RESULT 15
US-08-924-440-1
; Sequence 1, Application US/08924440
; Patent No. 5871550
;  GENERAL INFORMATION:
;    APPLICANT:  Frits et al.
;    TITLE OF INVENTION:  MUTANT THERMONOSPORA SPP. CELLULASE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International, Inc.
;      STREET:  925 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/924,440
;      FILING DATE:  August 27, 1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Christopher L. Stone
;      REGISTRATION NUMBER:  35,696
;      REFERENCE/DOCKET NUMBER:  GC388
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 846-7555
;      TELEFAX:  (650) 845-6504
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1293 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-924-440-1

  Query Match             4.9%;  Score 32.8;  DB 2;  Length 1293;
  Best Local Similarity   58.0%;  Pred. No. 2.1;
  Matches   58;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy     377 CTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGAC 436
           ||  ||  | || | ||  | ||| |  | ||| ||||       |||||  ||||| | |
Db    1080 CTACACCGGTGACGGCGCCAACGACTTCCAGATGGCCGACCGCTACATCGACCTGATGGC 1139

Qy     437 GGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
           ||| |  ||| |||  |||    |   || || |   ||||
Db    1140 GGAACGGAAGATCGGGTGGACCAAGTGGAACTACTCGGAC 1179


Search completed: December  9, 2002, 09:19:58
```

Job time : 57.9082 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:00:58 ; Search time 133.35 Seconds
(without alignments)
1962.092 Million cell updates/sec

Title: US-09-905-114-3
Perfect score: 671
Sequence: 1 acaacaggatctgccccata.....atcatcatgaagaaccatga 671

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	671	100.0	671	10	US-09-905-114-3	Sequence 3, Appli
2	592	88.2	592	10	US-09-905-114-1	Sequence 1, Appli
3	527	78.5	1023	10	US-09-962-436-336	Sequence 336, App
4	527	78.5	1023	10	US-09-954-456-271	Sequence 271, App
5	527	78.5	1023	10	US-09-880-107-3392	Sequence 3392, Ap
6	527	78.5	1079	12	US-10-133-065-1	Sequence 1, Appli
7	527	78.5	1079	12	US-10-074-509-1	Sequence 1, Appli
8	474.8	70.8	1124	12	US-10-133-065-11	Sequence 11, Appl
9	474.8	70.8	1124	12	US-10-074-509-11	Sequence 11, Appl
c 10	195.6	29.2	392	10	US-09-880-107-630	Sequence 630, App
11	163.2	24.3	1039	10	US-09-771-078-1	Sequence 1, Appli
12	163.2	24.3	1039	10	US-09-825-012-2	Sequence 2, Appli
13	161.6	24.1	783	10	US-09-825-012-4	Sequence 4, Appli
14	161.6	24.1	858	10	US-09-825-012-6	Sequence 6, Appli
15	161.6	24.1	1548	10	US-09-825-012-62	Sequence 62, Appl
16	161.6	24.1	1548	10	US-09-825-012-63	Sequence 63, Appl
17	161.6	24.1	1548	10	US-09-825-012-65	Sequence 65, Appl
18	161.6	24.1	1554	10	US-09-825-012-34	Sequence 34, Appl
19	161.6	24.1	1554	10	US-09-825-012-35	Sequence 35, Appl
20	161.6	24.1	1554	10	US-09-825-012-37	Sequence 37, Appl
21	161.6	24.1	1557	10	US-09-825-012-64	Sequence 64, Appl
22	161.6	24.1	1560	10	US-09-825-012-72	Sequence 72, Appl
23	161.6	24.1	1560	10	US-09-825-012-73	Sequence 73, Appl
24	161.6	24.1	1560	10	US-09-825-012-75	Sequence 75, Appl
25	161.6	24.1	1560	10	US-09-825-012-77	Sequence 77, Appl
26	161.6	24.1	1560	10	US-09-825-012-78	Sequence 78, Appl
27	161.6	24.1	1560	10	US-09-825-012-79	Sequence 79, Appl
28	161.6	24.1	1563	10	US-09-825-012-36	Sequence 36, Appl
29	161.6	24.1	1566	10	US-09-825-012-67	Sequence 67, Appl
30	161.6	24.1	1566	10	US-09-825-012-68	Sequence 68, Appl
31	161.6	24.1	1566	10	US-09-825-012-70	Sequence 70, Appl
32	161.6	24.1	1569	10	US-09-825-012-74	Sequence 74, Appl
33	161.6	24.1	1575	10	US-09-825-012-69	Sequence 69, Appl
34	161.6	24.1	1578	10	US-09-825-012-81	Sequence 81, Appl
35	161.6	24.1	1578	10	US-09-825-012-82	Sequence 82, Appl
36	161.6	24.1	1578	10	US-09-825-012-84	Sequence 84, Appl
37	161.6	24.1	1584	10	US-09-825-012-39	Sequence 39, Appl
38	161.6	24.1	1584	10	US-09-825-012-40	Sequence 40, Appl
39	161.6	24.1	1584	10	US-09-825-012-42	Sequence 42, Appl
40	161.6	24.1	1587	10	US-09-825-012-83	Sequence 83, Appl
41	161.6	24.1	1590	10	US-09-825-012-91	Sequence 91, Appl
42	161.6	24.1	1590	10	US-09-825-012-92	Sequence 92, Appl
43	161.6	24.1	1590	10	US-09-825-012-94	Sequence 94, Appl
44	161.6	24.1	1593	10	US-09-825-012-41	Sequence 41, Appl
45	161.6	24.1	1596	10	US-09-825-012-86	Sequence 86, Appl

ALIGNMENTS

```
RESULT 1
US-09-905-114-3
; Sequence 3, Application US/09905114
; Patent No. US20020048745A1
; GENERAL INFORMATION:
;  APPLICANT: ZHANG, HUANMIN
;  APPLICANT:  AX, ROY L
;  APPLICANT:  BELLIN, MARY E
;  TITLE OF INVENTION: ISOLATED POLYNUCLEOTIDE SEQUENCES ENCODING A FERTILITY ASSOCIATED
;  FILE REFERENCE: 210707US20
;  CURRENT APPLICATION NUMBER: US/09/905,114
;  CURRENT FILING DATE:  2001-07-14
;  PRIOR APPLICATION NUMBER: US 60/218,140
;  PRIOR FILING DATE: 2000-07-14
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 671
;   TYPE: DNA
;   ORGANISM: Bos sp.
US-09-905-114-3

  Query Match             100.0%;  Score 671;  DB 10;  Length 671;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-205;
  Matches  671;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
```

```
Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCATCATG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCATCATG 660

Qy    661 AAGAACCATGA 671
          ||||||||||||
Db    661 AAGAACCATGA 671
```

RESULT 2
US-09-905-114-1
; Sequence 1, Application US/09905114
; Patent No. US20020048745A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, HUANMIN
; APPLICANT: AX, ROY L
; APPLICANT: BELLIN, MARY E
; TITLE OF INVENTION: ISOLATED POLYNUCLEOTIDE SEQUENCES ENCODING A FERTILITY ASSOCIATED
; FILE REFERENCE: 210707US20
; CURRENT APPLICATION NUMBER: US/09/905,114
; CURRENT FILING DATE: 2001-07-14
; PRIOR APPLICATION NUMBER: US 60/218,140
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Bos sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(591)
; OTHER INFORMATION:
US-09-905-114-1

Query Match 88.2%; Score 592; DB 10; Length 592;
Best Local Similarity 100.0%; Pred. No. 3.1e-180;
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     27 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 86
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGAAGCTAAACGGAAATTCAAGAAAAGGCATAACATACAACTATGTGATTAGCTCTCGC 60

Qy     87 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTTGGAAGAAACACATATAAAGAACAGTATGCCTTTCTCTATAAAGAAAAGCTAGTGTCT 120

Qy    147 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAAAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG 180

Qy    207 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAACCCTTTGTGGTCTGGTTCCAGTCACCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240

Qy    267 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCCTGCACACCACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTAC 300

Qy    327 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACAGATGTGAAACGTCGCTGGAATGCAGAGAATTTCATTTTCATGGGTGACTTCAATGCT 360

Qy    387 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAG 420

Qy    447 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCC 480

Qy    507 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TATGACAGGATCGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAAC 540

Qy    567 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
```

RESULT 3
US-09-962-436-336
; Sequence 336, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 336
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-336

Query Match 78.5%; Score 527; DB 10; Length 1023;
Best Local Similarity 87.8%; Pred. No. 3.4e-159;
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          |||||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db    215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | |||||||||||||||||||||||| |||||||||||||||||||||||||| |||||||
Db    275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db    335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db    395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454
```

```
Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | ||||||||  ||||||||||||  ||||||||||||||||||||  ||||||||||||  |
Db    455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          ||||  ||||||  |||  |||  ||||||||  ||  |||||||   ||||||||  ||  |||||||
Db    515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db    575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          ||||||||  |||||||  |||||||  |||  ||||||||||||||||||||||||||||||
Db    635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          ||||  ||||||||||  |||||  ||  ||||||||||  ||||||||||||||  | ||  |
Db    695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  ||||  ||  |||  | ||||||   ||  |||||  ||||||||||||||| | ||
Db    755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||||||  |   | | |  ||
Db    815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869
```

RESULT 4
US-09-954-456-271
; Sequence 271, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 271
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-271

Query Match 78.5%; Score 527; DB 10; Length 1023;
Best Local Similarity 87.8%; Pred. No. 3.4e-159;
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy      1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
          |||||||||||||||||||||||||||||||||||| ||| ||||||||| | ||||||||
Db    215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy     61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
          | ||||||||||||||||||||||  ||||||||||||||||||||||||||  |||||||
Db    275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy    121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
          |||||||  ||  |||||||  ||||||||  ||     ||  ||  |  ||||||  ||||||||||
Db    335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy    181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
          |||||||||||||||||||||||||  ||||||||||||||||||  ||  |||  |||
Db    395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy    241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
          | ||||||||  ||||||||||||  ||||||||||||||||||  ||||||||||||  |
Db    455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy    301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
          ||||  ||||||  |||  |||  ||||||||  ||  |||||||   ||||||||  ||  |||||||
Db    515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy    361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
          ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db    575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy    421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
          ||||||||  |||||||  |||||||  |||  ||||||||||||||||||||||||||||||
Db    635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy    481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
          ||||  ||||||||||  |||||  ||  ||||||||||  ||||||||||||||  | ||  |
Db    695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy    541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
          |||  ||||  ||  |||  | ||||||   ||  |||||  ||||||||||||||| | ||
Db    755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy    601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||||  | |||||||||||||||||||||||||||||||||  |   | | |  ||
Db    815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869
```

RESULT 5
US-09-880-107-3392
; Sequence 3392, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3392
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U56814
US-09-880-107-3392

Query Match 78.5%; Score 527; DB 10; Length 1023;

```
  Best Local Similarity   87.8%;  Pred. No. 3.4e-159;
  Matches  575;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
           |||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db     215 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 274

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
           | ||||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db     275 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 334

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
           ||||||| || |||||||| |||||||| ||     || || | |||||| ||||||||||
Db     335 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 394

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
           |||||||||||||||||||||||||| ||||||||||||||||||| ||| |||||||||
Db     395 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 454

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
           | |||||||| |||||||||||| |||||||||||||||||||| ||||||||||||  |
Db     455 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 514

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
           |||| |||||| ||| ||| |||||||| || ||||||| |||||||| || |||||||
Db     515 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 574

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     575 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 634

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
           |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||
Db     635 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
           |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db     695 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 754

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
           |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db     755 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 814

Qy     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
           |||||  | |||||||||||||||||||||||||||||| |   | | |  ||
Db     815 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 869


RESULT 6
US-10-133-065-1
; Sequence 1, Application US/10133065
; Patent No. US20020123122A1
;   GENERAL INFORMATION:
;        APPLICANT: Baker, Kevin P.
;                   Baron, Will F.
;        TITLE OF INVENTION: HUMAN DNASE
;        NUMBER OF SEQUENCES: 11
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genentech, Inc.
;             STREET: 1 DNA Way
;             CITY: South San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94080
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: WinPatin (Genentech)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/133,065
;             FILING DATE: 26-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/643,520
;             FILING DATE: 22/08/2000
;             APPLICATION NUMBER: 08/794827
;             FILING DATE: 04-Feb-1997
;             APPLICATION NUMBER: 60/109796
;             FILING DATE: 05-Feb-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Evans, David W
;             REGISTRATION NUMBER: *See attached Limited Recognition under 37
;                                  C.F.R. 10.9(b)
;             REFERENCE/DOCKET NUMBER: P1000R1C3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650/225-1739
;             TELEFAX: 650/952-9881
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1079 base pairs
;             TYPE: Nucleic Acid
;             STRANDEDNESS: Single
;             TOPOLOGY: Linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-133-065-1

  Query Match             78.5%;  Score 527;  DB 12;  Length 1079;
  Best Local Similarity   87.8%;  Pred. No. 3.5e-159;
  Matches  575;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
           |||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db     261 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 320

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
           | ||||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db     321 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 380

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
           ||||||| || |||||||| |||||||| ||     || || | |||||| ||||||||||
Db     381 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 440

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
           |||||||||||||||||||||||||| ||||||||||||||||||| ||| |||||||||
Db     441 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 500

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
           | |||||||| |||||||||||| |||||||||||||||||||| ||||||||||||  |
Db     501 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 560

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
           |||| |||||| ||| ||| |||||||| || ||||||| |||||||| || |||||||
Db     561 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 620

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     621 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 680

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
           |||||||| ||||||| ||||||| ||| |||||||||||||||||||||||||||||||
Db     681 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 740

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
           |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db     741 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 800

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
           |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db     801 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 860

Qy     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
           |||||  | |||||||||||||||||||||||||||||| |   | | |  ||
Db     861 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 915
```

```
RESULT 7
US-10-074-509-1
; Sequence 1, Application US/10074509
; Patent No. US20020142437A1
;  GENERAL INFORMATION:
;       APPLICANT: Baker, Kevin P.
;                  Baron, Will F.
;       TITLE OF INVENTION: HUMAN DNASE
;       NUMBER OF SEQUENCES: 11
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genentech, Inc.
;            STREET: 1 DNA Way
;            CITY: South San Francisco
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94080
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: WinPatin (Genentech)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/074,509
;            FILING DATE: 11-Feb-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/643,520
;            FILING DATE: 22/08/2000
;            APPLICATION NUMBER: 08/794827
;            FILING DATE: 04-Feb-1997
;            APPLICATION NUMBER: 60/109796
;            FILING DATE: 05-Feb-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Johnston, Sean A.
;            REGISTRATION NUMBER: 35,910
;            REFERENCE/DOCKET NUMBER: P1000R1C2
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650/225-3562
;            TELEFAX: 650/952-9881
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1079 base pairs
;            TYPE: Nucleic Acid
;            STRANDEDNESS: Single
;            TOPOLOGY: Linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-074-509-1

  Query Match             78.5%;  Score 527;  DB 12;  Length 1079;
  Best Local Similarity   87.8%;  Pred. No. 3.5e-159;
  Matches  575;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy       1 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAA 60
           |||||||||||||||||||||||||||||||||| ||| |||||||||| | ||||||||
Db     261 ACAACAGGATCTGCCCCATACTGATGGAGAAGCTGAACAGAAATTCAAGGAGAGGCATAA 320

Qy      61 CATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCT 120
           | |||||||||||||||||||| |||||||||||||||||||||||||| |||||||
Db     321 CGTACAACTATGTGATTAGCTCTCGGCTTGGAAGAAACACATATAAAGAACAATATGCCT 380

Qy     121 TTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGG 180
           ||||||| || |||||||| |||||||| ||    || || | |||||| ||||||||||
Db     381 TTCTCTACAAGGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG 440

Qy     181 CTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACA 240
           |||||||||||||||||||||||||| |||||||||||||||||| || ||| |||
Db     441 ATGGAGACGCAGATGTGTTTTCCAGGGAGCCCTTTGTGGTCTGGTTCCAATCTCCCCACA 500

Qy     241 CCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAG 300
           | |||||||| |||||||||||| |||||||||||||||||| |||||||||||||  |
Db     501 CTGCTGTCAAAGACTTCGTGATTATCCCCCTGCACACCACCCCAGAGACATCCGTTAAGG 560

Qy     301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATT 360
           |||| |||||| ||| ||| |||||||| || |||||||  |||||||| || |||||||
Db     561 AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT 620

Qy     361 TCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGG 420
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     621 TCATTTTCATGGGTGACTTCAATGCCGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGA 680

Qy     421 ACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCA 480
           |||||||| ||||||| ||||||| ||| ||||||||||||||||||||||||||||||
Db     681 ACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 740

Qy     481 CGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTG 540
           |||| |||||||||| ||||| || |||||||||| |||||||||||||| | || |
Db     741 CGGTGAAGAAGAGCACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAGAAATCG 800

Qy     541 TCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600
           |||  |||| || |||  | ||||||   || ||||| ||||||||||||||| | ||
Db     801 TCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGACTTCCAGAAAGCTTACAAGCTGA 860

Qy     601 CTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
           |||||  | |||||||||||||||||||||||||||||| |   | | |  ||
Db     861 CTGAAGAGGAGGCCCTGGATGTCAGCGACCACTTTCCAGTTGAATTTAAACTACA 915


RESULT 8
US-10-133-065-11
; Sequence 11, Application US/10133065
; Patent No. US20020123122A1
;  GENERAL INFORMATION:
;       APPLICANT: Baker, Kevin P.
;                  Baron, Will F.
;       TITLE OF INVENTION: HUMAN DNASE
;       NUMBER OF SEQUENCES: 11
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genentech, Inc.
;            STREET: 1 DNA Way
;            CITY: South San Francisco
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94080
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: WinPatin (Genentech)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/133,065
;            FILING DATE: 26-Apr-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/643,520
;            FILING DATE: 22/08/2000
;            APPLICATION NUMBER: 08/794827
;            FILING DATE: 04-Feb-1997
;            APPLICATION NUMBER: 60/109796
;            FILING DATE: 05-Feb-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Evans, David W
;            REGISTRATION NUMBER: *See attached Limited Recognition under 37
;                                 C.F.R. 10.9(b)
;            REFERENCE/DOCKET NUMBER: P1000R1C3
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650/225-1739
;            TELEFAX: 650/952-9881
;  INFORMATION FOR SEQ ID NO: 11:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1124 base pairs
;            TYPE: Nucleic Acid
;            STRANDEDNESS: Single
;            TOPOLOGY: Linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-133-065-11
```

```
  Query Match             70.8%;  Score 474.8;  DB 12;  Length 1124;
  Best Local Similarity   82.9%;  Pred. No. 2.1e-142;
  Matches  542;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy       2 CAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAAC 61
          |||||  |||||  |||||  ||||||||||||||  ||  |||||||||  |||  |  |||  |||
Db     379 CAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAATTCACGAAGAAGCACAAC 438

Qy      62 ATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCTT 121
          ||||||||||||||||||  |||||  ||||||||||||||  ||  |||||  ||||||||||
Db     439 ATACAACTATGTGATTAGTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTATGCCTT 498

Qy     122 TCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGC 181
            ||||  ||  ||  |||||  |||||||  ||    ||   ||||  ||||||  |||||||||
Db     499 CGTCTACAAGGAGAAGCTGGTGTCTGTGAAGACAAAATACCACTACCATGACTATCAGGA 558

Qy     182 TGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACAC 241
          |||||||  ||||  ||||||||||||  ||||||||||  ||||||||  ||  ||||   ||
Db     559 TGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTAC 618

Qy     242 CGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGA 301
          |||||||||||||||||||||||||||  |||||||  ||  ||  |||||  |||||||  |||
Db     619 TGCTGTCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTTAAAGA 678

Qy     302 GATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTT 361
          |||  ||||||||||  ||||||||||  |||||||  |  |  |  |||||  |||||||||
Db     679 GATAGATGAGCTGGTCGATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTT 738

Qy     362 CATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGA 421
          |||  ||||||||||  |||||  ||  |||||  |||||  ||||||||||||||||||  ||  |
Db     739 CATCTTCATGGGTGATTTCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAA 798

Qy     422 CATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCAC 481
          |||  ||  ||||||||||||||||||  ||||||||||  ||||||||||||||||||  ||
Db     799 CATTCGTTTGAGGACGGACCCCAAGTTTGTTTGGCTGATTGGGGACCAAGAGGACACTAC 858

Qy     482 GGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGT 541
          ||||||||||||  ||  |  |||  ||||||||||||||  ||||||  |  ||||||  |  ||  ||
Db     859 GGTCAAGAAGAGTACCAGCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAGATAGT 918

Qy     542 CAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTC 601
          ||||||  |  |  |||  |  ||  |    |||||||||  ||  ||||||||||    |||||
Db     919 CAACTCCGTGGTTCCCCGTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTC 978

Qy     602 TGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||   | ||||||||||||||||  ||  ||||||||||||  |   |  |  ||
Db     979 TGAGGAGGAGGCCCTGGATGTCAGTGATCACTTTCCAGTTGAGTTTAAGCTACA 1032


RESULT 9
US-10-074-509-11
; Sequence 11, Application US/10074509
; Patent No. US20020142437A1
;   GENERAL INFORMATION:
;        APPLICANT: Baker, Kevin P.
;                   Baron, Will F.
;        TITLE OF INVENTION: HUMAN DNASE
;        NUMBER OF SEQUENCES: 11
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genentech, Inc.
;             STREET: 1 DNA Way
;             CITY: South San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94080
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: WinPatin (Genentech)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/074,509
;             FILING DATE: 11-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/643,520
;             FILING DATE: 22/08/2000
;             APPLICATION NUMBER: 08/794827
;             FILING DATE: 04-Feb-1997
;             APPLICATION NUMBER: 60/109796
;             FILING DATE: 05-Feb-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Johnston, Sean A.
;             REGISTRATION NUMBER: 35,910
;             REFERENCE/DOCKET NUMBER: P1000R1C2
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650/225-3562
;             TELEFAX: 650/952-9881
;   INFORMATION FOR SEQ ID NO: 11:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1124 base pairs
;             TYPE: Nucleic Acid
;             STRANDEDNESS: Single
;             TOPOLOGY: Linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-074-509-11

  Query Match             70.8%;  Score 474.8;  DB 12;  Length 1124;
  Best Local Similarity   82.9%;  Pred. No. 2.1e-142;
  Matches  542;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy       2 CAACAGGATCTGCCCCATACTGATGGAGAAGCTAAACGGAAATTCAAGAAAAGGCATAAC 61
          |||||  |||||  |||||  ||||||||||||||  ||  |||||||||  |||  |  |||  |||
Db     379 CAACAACATCTGTCCCATGCTGATGGAGAAGCTGAATGGAAATTCACGAAGAAGCACAAC 438

Qy      62 ATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCCTT 121
          ||||||||||||||||||  |||||  ||||||||||||||  ||  |||||  ||||||||||
Db     439 ATACAACTATGTGATTAGTTCTCGACTTGGAAGAAACACGTACAAAGAGCAGTATGCCTT 498

Qy     122 TCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGACTATCAGGC 181
            ||||  ||  ||  |||||  |||||||  ||    ||   ||||  ||||||  |||||||||
Db     499 CGTCTACAAGGAGAAGCTGGTGTCTGTGAAGACAAAATACCACTACCATGACTATCAGGA 558

Qy     182 TGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCCTACAC 241
          |||||||  ||||  ||||||||||||  ||||||||||  ||||||||  ||  ||||   ||
Db     559 TGGAGACACAGACGTGTTTTCCAGGGAGCCCTTTGTGGTTTGGTTCCATTCCCCCTTTAC 618

Qy     242 CGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTTAGAGA 301
          |||||||||||||||||||||||||||  |||||||  ||  ||  |||||  |||||||  |||
Db     619 TGCTGTCAAGGACTTCGTGATTGTCCCCTTGCACACAACTCCCGAGACCTCCGTTAAAGA 678

Qy     302 GATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAATTT 361
          |||  ||||||||||  ||||||||||  |||||||  |  |  |  |||||  |||||||||
Db     679 GATAGATGAGCTGGTCGATGTCTACACGGATGTGAGAAGCCAGTGGAAGACAGAGAATTT 738

Qy     362 CATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGGAAGGA 421
          |||  ||||||||||  |||||  ||  |||||  |||||  ||||||||||||||||||  ||  |
Db     739 CATCTTCATGGGTGATTTCAACGCCGGCTGTAGCTATGTCCCCAAGAAGGCCTGGCAGAA 798

Qy     422 CATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGACACCAC 481
          |||  ||  ||||||||||||||||||  ||||||||||  ||||||||||||||||||  ||
Db     799 CATTCGTTTGAGGACGGACCCCAAGTTTGTTTGGCTGATTGGGGACCAAGAGGACACTAC 858

Qy     482 GGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAATATTGT 541
          ||||||||||||  ||  |  |||  ||||||||||||||  ||||||  |  ||||||  |  ||  ||
Db     859 GGTCAAGAAGAGTACCAGCTGTGCCTATGACAGGATTGTGCTTTGTGGACAAGAGATAGT 918

Qy     542 CAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTC 601
          ||||||  |  |  |||  |  ||  |    |||||||||  ||  ||||||||||    |||||
Db     919 CAACTCCGTGGTTCCCCGTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATGACTTGTC 978

Qy     602 TGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGTTCATCATCATCATCA 655
          |||   | ||||||||||||||||  ||  ||||||||||||  |   |  |  ||
Db     979 TGAGGAGGAGGCCCTGGATGTCAGTGATCACTTTCCAGTTGAGTTTAAGCTACA 1032
```

```
RESULT 10
US-09-880-107-630/c
; Sequence 630, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 630
;   LENGTH: 392
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA279676
US-09-880-107-630

  Query Match             29.2%;  Score 195.6;  DB 10;  Length 392;
  Best Local Similarity   84.9%;  Pred. No. 5.4e-53;
  Matches  219;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy     398 CGTCCCCAAGAAGGCCTGGAAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCT 457
           ||||||||||||||||||||||| |||||||| ||||||| ||||||| ||| ||||||||
Db     392 CGTCCCCAAGAAGGCCTGGAAGAACATCCGCTTGAGGACTGACCCCAGGTTTGTTTGGCT 333

Qy     458 GATCGGGGACCAAGAGGACACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGAT 517
           |||||||||||||||||||||||||| |||||||||| ||||| || ||||||||||
Db     332 GATCGGGGACCAAGAGGACACCACGGTGAAGAAGAGCACCAACTGTGCATATGACAGGAT 273

Qy     518 CGTGCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGA 577
           |||||||||||||||| | || |||| |||| || ||| | |||||| || |||||
Db     272 TGTGCTTAGAGGACAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAACAGTGTTTTTGA 213

Qy     578 TTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCC 637
           |||||||||||||||| | || |||||  | ||||||||||||||||||||||||||||
Db     212 CTTCCAGAAAGCTTACAAGCTGACTGAAAAGGAGGCCCTGGATGTCAGCGACCACTTTCC 153

Qy     638 AGTTCATCATCATCATCA 655
           |||| |   | | |  ||
Db     152 AGTTGAATTTAAACTACA 135


RESULT 11
US-09-771-078-1
; Sequence 1, Application US/09771078
; Patent No. US20020034727A1
;  GENERAL INFORMATION:
;    APPLICANT:  Mrsny, Randall J.,
;    APPLICANT:  Daugherty, Ann L.,
;    APPLICANT:  Patapoff, Thomas W.
;    TITLE OF INVENTION:  DNase Compaction Assay
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/771,078
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/971,019
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Glaister, Debra J.
;      REGISTRATION NUMBER:  33,888
;      REFERENCE/DOCKET NUMBER:  792
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3177
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1039 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-771-078-1

  Query Match             24.3%;  Score 163.2;  DB 10;  Length 1039;
  Best Local Similarity   57.4%;  Pred. No. 2.7e-42;
  Matches  335;  Conservative    0;  Mismatches  243;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 176
           ||  | || |       | |  |||||||       | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
           || |  | |     ||    ||   | | || ||     | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
           | ||| |  |||| ||| || |  ||||| |||||||  |  |||| | |    | ||
Db     589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
           |||||| || |  ||   ||| ||||||   |||||  ||      |||      |||
Db     649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
           |  |||| || ||||| |||||||||| ||||||||||  ||              |||
Db     709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
           |||||||||| ||||   |||||  |||   |||||||||   |||   |  |||
Db     769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
           ||||| |      |   |||  |||| ||||||||||||||||||  |||
Db     829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy     537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
           |      | | | || ||| | ||     ||   |||| | ||||||  || ||  |
Db     886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy     597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
           ||  |||      |||||  |   |||| ||||||| ||||||
Db     946 CTGAGTGACCAACTGGCCCAAGCCATCAGTGACCACTATCCAGT 989


RESULT 12
US-09-825-012-2
; Sequence 2, Application US/09825012
; Patent No. US20020122798A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Young, Robert
;  TITLE OF INVENTION: Compounds for Targeting
;  FILE REFERENCE: 43191-256808
;  CURRENT APPLICATION NUMBER: US/09/825,012
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/237,159
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: GB 0008049.9
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1039
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-012-2

  Query Match             24.3%;  Score 163.2;  DB 10;  Length 1039;
  Best Local Similarity   57.4%;  Pred. No. 2.7e-42;
  Matches  335;  Conservative    0;  Mismatches  243;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     409 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 176
           ||  | || |      | |  |||||||       | ||||||  |||| | ||   ||
Db     469 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 528

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
           || |  | |      ||    ||   | | || ||     | ||| |||||   || |
Db     529 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 588

Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
           | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |    | ||
Db     589 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 648

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
             ||||| || |  ||   ||| ||||||   |||||  ||       |||      |||
Db     649 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
           |  |||| || ||||| |||||||||| |||||||||| ||              |||
Db     709 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 768

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
           |||||||||| ||||  |||||  |||  |||||||||  |||  |  |||
Db     769 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
           |||||  |     |    |||  |||| ||||||||||||||||||  ||  |||
Db     829 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 885

Qy     537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
           |       | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db     886 CTCCGAGGCGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 945

Qy     597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
           ||  |||      |||||  |   |||| ||||||| ||||||
Db     946 CTGAGTGACCAACTGGCCCAAGCCATCAGTGACCACTATCCAGT 989


RESULT 13
US-09-825-012-4
; Sequence 4, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Robert
;  TITLE OF INVENTION: Compounds for Targeting
;  FILE REFERENCE: 43191-256808
;  CURRENT APPLICATION NUMBER: US/09/825,012
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/237,159
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: GB 0008049.9
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 783
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-012-4

  Query Match             24.1%;  Score 161.6;  DB 10;  Length 783;
  Best Local Similarity   57.2%;  Pred. No. 7.2e-42;
  Matches  334;  Conservative    0;  Mismatches  244;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     184 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 243

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 176
           ||  | || |      | |  |||||||       | ||||||  |||| | ||   ||
Db     244 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 303

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
           || |  | |      ||    ||   | | || ||     | ||| |||||   || |
Db     304 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 363

Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
           | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |    | ||
Db     364 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 423

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
           |||||  || |  ||   ||| ||||||   |||||  ||       |||      |||
Db     424 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 483

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
           |  |||| || ||||| |||||||||| |||||||||| ||              |||
Db     484 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 543

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
           |||||||||| ||||  |||||  |||  |||||||||  |||  |  |||
Db     544 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 603

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
           |||||  |     |    |||  |||| ||||||||||||||||||  ||  |||
Db     604 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 660

Qy     537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
           |       | | | || |||  | ||    ||   |||| | ||||||   || ||  |
Db     661 CTCCGAGGGGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 720

Qy     597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
           ||  |||      |||||  |   |||| ||||||| ||||||
Db     721 CTGAGTGACCAACTGGCCCAAGCCATCAGTGACCACTATCCAGT 764


RESULT 14
US-09-825-012-6
; Sequence 6, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Robert
;  TITLE OF INVENTION: Compounds for Targeting
;  FILE REFERENCE: 43191-256808
;  CURRENT APPLICATION NUMBER: US/09/825,012
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/237,159
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: GB 0008049.9
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 6
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-012-6

  Query Match           24.1%;  Score 161.6;  DB 10;  Length 858;
  Best Local Similarity   57.2%;  Pred. No. 7.6e-42;
  Matches  334;  Conservative    0;  Mismatches  244;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     259 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 318

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 176
           ||  | || |      | |  |||||||      | ||||||  |||| | ||   ||
Db     319 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 378

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
           || |  | |      ||    ||   | | || ||     | ||| |||||   || |
Db     379 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 438

Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
           | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |    | ||
Db     439 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 498

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
           |||||  || |  ||   ||| ||||||   |||||  ||      |||      |||
Db     499 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 558

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
           |  |||| || ||||| |||||||||| ||||||||||  ||          |||
Db     559 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 618

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
           ||||||||||| ||||  |||||  |||  |||||||||  |||  |  |||
Db     619 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 678

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
           |||||  |     |   |||  |||| |||||||||||||||||  ||  |||
Db     679 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 735

Qy     537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
           |     | | || |||  | ||   ||  ||||  | ||||||   || ||  |
Db     736 CTCCGAGGGGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 795

Qy     597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
           ||  |||      |||||  |   |||| ||||||| ||||||
Db     796 CTGAGTGACCAACTGGCCCAAGCCATCAGTGACCACTATCCAGT 839


RESULT 15
US-09-825-012-62
; Sequence 62, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Robert
;  TITLE OF INVENTION: Compounds for Targeting
;  FILE REFERENCE: 43191-256808
;  CURRENT APPLICATION NUMBER: US/09/825,012
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/237,159
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: GB 0008049.9
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
;   LENGTH: 1548
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Humanised HMFG1 Fd - DNase I fusion

US-09-825-012-62

  Query Match           24.1%;  Score 161.6;  DB 10;  Length 1548;
  Best Local Similarity   57.2%;  Pred. No. 1.1e-41;
  Matches  334;  Conservative    0;  Mismatches  244;  Indels    6;  Gaps    2;

Qy      60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAAGAAACACATATAAAGAACAGTATGCC 119
           || ||  |||| ||| | ||    |  || ||| | ||||  ||||| || |  ||
Db     949 ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 1008

Qy     120 TTTCTCTATAAAGAAAAGCTAGTGTCTGTAAAACAAAGCTACCTCTACCACGA---CTAT 176
           ||  | || |      | |  |||||||      | ||||||  |||| | ||   ||
Db    1009 TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGGCTGC 1068

Qy     177 CAGGCTGGAGACGCAGATGTGTTTTCCAGGGAACCCTTTGTGGTCTGGTTCCAGTCACCC 236
           || |  | |      ||    ||   | | || ||     | ||| |||||   || |
Db    1069 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG 1128

Qy     237 TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCACCCCTGAGACATCCGTT 296
           | ||| |  |||| ||| || |  ||||| ||||||||  |  |||| | |    | ||
Db    1129 TTCACAGAGGTCAGGGAGTTTGCCATTGTTCCCCTGCATGCGGCCCCGGGGGACGCAGTA 1188

Qy     297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGCAGAG 356
           |||||  || |  ||   ||| ||||||   |||||  ||      |||      |||
Db    1189 GCCGAGATCGACGCTCTCTATGACGTCTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 1248

Qy     357 AATTTCATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCCAAGAAGGCCTGG 416
           |  |||| || ||||| |||||||||| ||||||||||  ||          |||
Db    1249 GACGTCATGTTGATGGGCGACTTCAATGCGGGCTGCAGCTATGTGAGACCCTCCCAGTGG 1308

Qy     417 AAGGACATCCGCCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
           ||||||||||| ||||  |||||  |||  |||||||||  |||  |  |||
Db    1309 TCATCCATCCGCCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 1368

Qy     477 ACCACGGTCAAGAAGAGCACAAACTGCGCCTATGACAGGATCGTGCTTAGAGGACAAAAT 536
           |||||  |     |   |||  |||| |||||||||||||||||  ||  |||
Db    1369 ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGGATGCTG 1425

Qy     537 ATTGTCAACTCTGGTGGTCCTCAATCAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
           |     | | || |||  | ||   ||  ||||  | ||||||   || ||  |
Db    1426 CTCCGAGGGGCCGTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGGC 1485

Qy     597 TTGTCTGAATCGAAGGCCCTGGATGTCAGCGACCACTTTCCAGT 640
           ||  |||      |||||  |   |||| ||||||| ||||||
Db    1486 CTGAGTGACCAACTGGCCCAAGCCATCAGTGACCACTATCCAGT 1529


Search completed: December  9, 2002, 09:24:18
Job time : 144.35 secs
```